



Etude du rôle de dux4 dans la physiopathologie de la dystrophie facio-scapulo-humérale

Maxime Ferreboeuf

► To cite this version:

Maxime Ferreboeuf. Etude du rôle de dux4 dans la physiopathologie de la dystrophie facio-scapulo-humérale. Biologie moléculaire. Université Pierre et Marie Curie - Paris VI, 2014. Français. NNT : 2014PA066267 . tel-01340357

HAL Id: tel-01340357

<https://theses.hal.science/tel-01340357>

Submitted on 1 Jul 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Pierre et Marie Curie

Ecole doctorale Complexité du vivant

Institut de Myologie - Centre de Recherche

Titre de la thèse :

Etude du rôle de DUX4 dans la physiopathologie de la Dystrophie Facio-Scapulo-Humérale

Par Mr FERREBOEUF Maxime

Thèse de doctorat de Biologie Moléculaire et Cellulaire

Dirigée par DUMONCEAUX Julie

Présentée et soutenue publiquement le 30 juin 2014

Devant un jury composé de :

Monsieur CHARLET-BERGUERAND Nicolas, Rapporteur

Madame MARTINAT Cécile, Rapporteur

Madame MAGDINIER Frédérique, Examineur

Monsieur BASSEZ Guillaume, Examineur

Monsieur HUBE Florent, Examineur

Monsieur AMSELEM Serge, Président du Jury

Madame DUMONCEAUX Julie, Directrice de thèse

RESUME

La Dystrophie Facio-Scapulo-Humérale (FSHD) est la plus fréquente maladie neuromusculaire chez l'adulte (7 cas pour 100 000 naissances). D'origine autosomique dominante, elle apparaît en général durant la seconde décade et se caractérise principalement par une perte de force et une fonte asymétrique des muscles du visage, des épaules et des bras puis progresse au reste du corps. Cette pathologie est liée dans 95% des cas (FSHD 1) à une délétion d'un motif *D4Z4* de 3,3Kb répété en tandem à l'extrémité subtelomérique du chromosome 4. Chez les individus non malades, le nombre de répétition est compris entre 11 et 150 copies alors que les patients FSHD 1 présentent un nombre réduit compris entre 1 à 10. Un second groupe de patients (FSHD 2), représentant environ 5% des cas, possède un nombre de répétitions normal mais ne présente aucune différence phénotypique avec les patients FSHD 1. Une hypométhylation commune aux patients FSHD 1 et 2 et l'absence de différences phénotypiques suggèrent un mécanisme physiopathologique commun. L'analyse du motif *D4Z4* a mis en évidence la présence d'une phase ouverte de lecture codant pour un facteur de transcription appelé DUX4. Son expression a été retrouvée spécifiquement chez les deux groupes de patients FSHD et un grand nombre de gènes dérégulés par DUX4 ont été mis en évidence, mais aucun lien n'a été établi entre l'expression de DUX4, les dérégulations géniques et le développement de la pathologie. Mes travaux de thèse se sont concentrés sur l'étude de DUX4 et son implication dans la physiopathologie de la FSHD. Parce que la FSHD est une maladie progressive, nos analyses ont été faites sur des cultures de cellules et des biopsies fœtales FSHD 1 et Contrôles. Nous avons ainsi mis en évidence pour la première fois que DUX4 est exprimé dès le stade fœtal et que les dérégulations géniques identifiées chez l'adulte sont également présentes à ce stade. De plus nos expériences semblent montrer que le niveau d'expression de *DUX4* entre le fœtus et l'adulte FSHD est similaire. Le niveau d'expression de DUX4 est extrêmement restreint (environ 0,5 à 10% des noyaux), mais son effet sur un grand nombre de gènes cibles est lui très fortement prononcé et quantifiable. C'est pourquoi il a été postulé que DUX4 serait à même de diffuser d'un noyau à l'autre au sein d'une même fibre musculaire et entraîner des dérégulations géniques en cascades. Des essais de co-cultures entre des cellules Humaine FSHD et Murine suivis d'immunomarquage et de quantification de gènes murins, activés par DUX4, ont confirmé notre hypothèse suggérant ainsi que *in vivo*, DUX4 est à même de diffuser le phénomène pathologique.

Nos travaux ouvrent ainsi un nouveau champ d'étude pour la compréhension des mécanismes physiopathologiques conduisant à la FSHD.

Remerciements

Je tiens tout d'abord à remercier tous les membres du jury pour avoir accepté de lire et de juger ce travail.

Je remercie la FSHD Global Research Foundation pour avoir financé mes trois années de thèse.

Je souhaite remercier tout particulièrement Julie DUMONCEAUX pour m'avoir encadré durant ces trois années. Je te remercie de m'avoir transmis tes connaissances techniques avec parfois, je dois le dire, un certain étonnement de ma part (mais Julie pourquoi tu rajoutes de l'enzyme pendant que le thermocycleur tourne ? et bien je ne sais pas, je fais des tests !! ah d'accord). Je te remercie d'avoir cru en moi et d'avoir tout fait afin d'obtenir un financement pour ces trois années de thèse. Je te remercie également pour la grande disponibilité dont tu as fait preuve jusqu'à la fin de ce travail. Ces trois années ont été riches en émotions, et malgré certaines périodes difficiles, ta franchise et ton enthousiasme m'ont permis de me construire scientifiquement et personnellement. Pour tout cela, je souhaiterais t'exprimer ma profonde reconnaissance.

Je souhaite également remercier l'ensemble des membres de la Team FSH, les anciens et les nouveaux, et tout particulièrement Virginie pour m'avoir permis de mener mes projets scientifiques à bien.

J'aimerais remercier Claire Francastel et Florent pour ces discussions scientifiques enrichissantes et leur bonne humeur.

Je voudrais remercier Thomas Voit, directeur de l'unité de recherche U974, pour m'avoir accueilli dans son laboratoire.

Je souhaiterais également remercier Gillian et Vincent, pour m'avoir accueilli dans la Team 3 (et ex Team 2) ainsi que pour leur sympathie de tous les jours et le partage de leurs connaissances scientifiques durant nos réunions.

J'aimerais remercier chaque membre de l'équipe DM qui ont su être présent lorsque j'en avais besoin :

Arnaud : un grand merci pour tes conseils avisés et ta grande disponibilité. Tu as su m'écouter lorsque j'en avais besoin et m'aider à relever la tête. Nous avons partagé un grand nombre de fous rires et de débats scientifiques et tu auras réussi au moins une chose, me faire écouter autre chose que de la « techno » et cela ce n'est pas une petite victoire de ta part.

Frédo : Je te considère comme ma grande sœur (cela tombe bien vous avez le même âge !!), je suis très heureux de te connaître et tes connaissances scientifiques m'ont beaucoup apporté. Que de fous rires partagés (dont certains douteux : plus jamais...) et tout comme Arnaud, tu m'as aidé dans les moments difficiles, un grand merci. Tu étais ma seule alliée pour le choix de la station radio dans notre laboratoire et pour cela je te remercie également. Enfin tu m'a également appris quelques chose : « Y'a quelque chose besoin ? ».

Coralie (allias GC) : Ma collègue de bibliothèque des premiers mois... nous sommes arrivés à la même période au laboratoire et nous revoilà presque 4 ans plus tard. Nous avons partagé d'excellents moments, je te remercie pour ton amitié et te souhaite bon courage pour ta soutenance prochaine.

Aux autres membres DM: Reymond, Arnaud J (j'suis un barbare, aucune différence...), Joëlle et Tonton (non pas la claque !!!), merci beaucoup pour votre bonne humeur au quotidien, qui a contribué à mon épanouissement dans ce laboratoire.

J'aimerais également remercier les autres membres du laboratoire dont l'aide et la bonne humeur quotidienne ont participé à mon intégration et au bon déroulement de ces trois années.

Une thèse est une expérience unique avec des périodes agréables suivies de moments stressants. C'est pourquoi j'aimerais terminer par remercier mes proches qui ont été présents et m'ont soutenu durant toutes ces années.

Jessica, je te remercie pour chaque chose que tu m'apportes au quotidien, ta présence est un vrai bonheur et est pour moi un point de repère dans ma vie.

Florence, un immense merci pour ton aide lors de cette rédaction, tes relectures attentives ont contribué à la qualité de ce manuscrit.

Maman, Papa, j'aimerais vous remercier pour m'avoir soutenu, sans vous, je ne serais pas où j'en suis aujourd'hui. Un grand merci pour votre présence au quotidien, je vous dédie ce travail.

TABLE DES MATIERES

TABLE DES MATIERES	6
LISTE DES FIGURES.....	11
INTRODUCTION.....	15
1. Historique.....	16
2. Physiopathologie de la dystrophie Facio-Scapulo-Humérale	16
2.1 Une atteinte du muscle squelettique.....	16
2.2 Atteintes musculaires des patients FSHD	18
2.3 Atteintes non musculaires	20
3. Aspect génomique de la FSHD	20
3.1 Identification d'un motif subtélomérique.....	20
3.2 Le motif <i>D4Z4</i> est également présent sur le chromosome 10	22
3.3 Patients mosaïques	22
3.4 Diagnostic génétique par Southern blot	24
3.5 Complexité de la FSHD	25
3.5.1 Identification de différents segments polymorphes	25
3.5.2 Prévalence des haplotypes.....	27
3.6 Diagnostic par peignage moléculaire	28
4. Modèles de régulations épigénétiques.....	29
4.1 Méthylation des régions <i>D4Z4</i>	29
4.2 L'hypométhylation est également retrouvée chez les patients FSHD 2.....	31
4.3 Implication des télomères (TPE) et effets de positions.....	32
4.4 Domaine d'attache S/MAR et localisation nucléaire des répétitions <i>D4Z4</i>	34
4.5 Identification de Domaines présents dans <i>D4Z4</i>	36
4.5.1 Domaine Enhancer	36
4.5.2 Domaine Insulateur <i>D4Z4</i>	38
4.5.3 Domaine répresseur DBE.....	39
4.6 Capacité bidirectionnelle du promoteur <i>D4Z4</i>	42
5. Gènes dérégulés dans la FSHD	43
5.1 FRG1	43
5.2 FRG2	44
5.3 ANT1.....	45

5.4	FAT1	45
6.	Le modèle candidat : DUX4.....	47
6.1	Le motif D4Z4.....	47
6.2	DUX4 : un facteur de transcription	49
6.2.1	Les différents transcrits identifiés	49
6.2.2	Expression chez les patients FSHD.....	50
6.2.3	Un modèle d'expression pour DUX4.....	52
6.3	Conséquences liées à l'expression de DUX4	53
6.3.1	Toxicité et dérégulations <i>in vitro</i>	53
6.3.2	Toxicité et dérégulations <i>in vivo</i>	55
	RESULTATS.....	59
	OBJECTIFS DE LA THESE	60
	PUBLICATION N°1: DUX4 and DUX4 downstream target genes are expressed in fetal FSHD muscles.....	61
	PUBLICATION N°2: Nuclear protein spreading: implication for pathophysiology of neuromuscular diseases	74
	AUTRES RESULTATS : Un micro ARN issu de l'intron 1 de DUX4 ?	84
1.	Introduction	85
1.1	Les micros ARN.....	85
1.2	Fonction des micros ARN	85
2.	Biogenèse des micros ARN.....	86
2.1	Transcription des miARN	86
2.2	Clivage par la machinerie DROSHA-DGCR8 et export.....	87
2.3	Maturation cytoplasmique médiée par DICER	88
3.	Les micros ARN introniques.....	89
4.	Cas particulier des mirtrons.....	90
5.	Résultats	91
5.1	Identification d'un petit ARN non codant par bioinformatique	91
5.2	Mise en évidence d'un ARN non codant par Northern Blot	93
6.	Discussion et perspectives.....	95
6.1	Résultats à confirmer.....	95
6.2	Caractérisation de la voie de production du micro ARN	95
6.3	Vérification de la fonctionnalité du miARN	96

6.4	Le miARN peut-il venir réguler DUX4 ?	97
7.	Matériels et méthodes	98
7.1	Extraction et séparation des ARN	98
7.2	Préparation et migration sur gel de polyacrylamide	98
7.3	Transfert sur membrane de nitrocellulose et cross link chimique	99
7.4	Marquage de la sonde oligonucléotide, Hybridation et Révélation	99
	DISCUSSION ET PERSPECTIVES	100
1.	DUX4, un lien commun unifiant les patients FSHD1 et FSHD2	101
2.	Une fonction pour les différentes isoformes de <i>DUX4</i> ?	101
2.1	Récapitulatif des différentes isoformes de <i>DUX4</i>	101
2.2	La production des isoformes <i>DUX4-S</i> et <i>DUX4-FL</i> régulée par le statut épigénétique de la cellule ?	102
2.3	Des isoformes alternatives de <i>DUX4</i> présentes dans le muscle squelettique des patients FSHD	103
2.4	DUX4 : un facteur de transcription nécessaire dans les gonades mais toxique dans le muscle ?	104
2.5	DUX4, un gène, plusieurs protéines ?	105
3.	Expression de DUX4 dans des cellules et biopsies FSHD	107
3.1	Cinétique d'expression <i>in vitro</i> dans des cellules fœtales FSHD	107
3.2	DUX4 est également détecté <i>in vivo</i> dans des biopsies fœtales et adultes FSHD	108
3.3	Une détection de DUX4 plus difficile dans les biopsies	108
4.	L'expression de DUX4 est-elle liée aux vagues de myogenèse ?	109
5.	Dérégulations de gènes impliqués dans la myogénèse	110
6.	Altération de l'expression génique en cis et en trans chez le fœtus porteur de contractions <i>D4Z4</i>	112
6.1	Gènes du locus 4q35	113
6.2	Dérégulations géniques en lien avec DUX4 chez le fœtus et l'adulte FSHD	115
7.	Comment cette protéine faiblement exprimée peut induire des dérégulations géniques majeures ?	115
8.	Un autre gène retrouvé dérégulé chez le fœtus porteur de contractions <i>D4Z4</i> : <i>FAT1</i>	117
8.1	Un lien entre DUX4 et FAT 1 ?	118
9.	Quelles origines pour expliquer les atteintes musculaires différentielles ?	119
9.1	Origine somitique ?	119

9.2	Différences intrinsèques entre les muscles ?	120
10.	Enfin, DUX4 a-t-il un rôle dans la pathologie ?	121
	CONCLUSION GENERALE	123
	BIBLIOGRAPHIES.....	125
	ANNEXES.....	141
	PUBLICATION N°3: Dysregulation of 4q35- and muscle-specific gene in fetuses with a short D4Z4 array linked to facio-scapulo-humeral dystrophy	142

LISTE DES FIGURES

	Page
Figure 1 : Régénération du muscle squelettique	16
Figure 2 : Planche anatomique et photos de patients récapitulant la répartition des principaux muscles touchés ainsi que les signes caractéristiques de la dystrophie Facio-Scapulo-Humérale	18
Figure 3 : Motif <i>D4Z4</i> sur le bras long du chromosome 4 en 4q35 chez les individus sains et atteints de FSHD	20
Figure 4 : Génération et immortalisation de clones cellulaires provenant de patients mosaïques	23
Figure 5 : Répartition des enzymes de restrictions sur les répétitions <i>D4Z4</i> entre le chromosome 4 et 10 ainsi que de la sonde P13E-11 utilisée pour mesurer la taille des répétitions par Southern Blot	24
Figure 6 : Organisation des polymorphismes sur les répétitions <i>D4Z4</i> des chromosomes 4 et 10	25
Figure 7 : Répartition des haplotypes au sein de 223 patients FSHD	26
Figure 8 : Technique de peignage moléculaire	28
Figure 9 : Modèle schématique des changements chromatinien liés à la perte des H3K9me3 chez les patients FSHD	29
Figure 10 : Récapitulation des différences entre les patients FSHD 1 et FSHD 2	31
Figure 11 : Modèle de l'effet de position télomérique (TPE) sur la régulation des gènes du locus 4q35	32
Figure 12 : Modèle d'attache à la membrane nucléaire et des boucles chromosomiques	33
Figure 13 : Corrélation entre orientation spatiale et densité génique du chromosome 4 et 17	34
Figure 14 : Répartition des domaines présents dans les répétitions <i>D4Z4</i>	35
Figure 15 : Modèle schématique de l'activation de l'enhancer dans les répétitions <i>D4Z4</i> par le facteur KLF15	36
Figure 16 : Modèle expliquant le rôle de l'insulateur <i>D4Z4</i> et son implication dans les modifications épigénétiques de la FSHD	38

Figure 17 : Modèle de répression transcriptionnelle médiée par la fixation du complexe multiprotéique YY1, HMGB1 et Nucleoline sur le domaine DBE	39
Figure 18 : Modèle de régulation épigénétique du locus 4q35 par les complexes polycombs/trithorax	40
Figure 19 : Représentation des gènes étudiés dans la FSHD sur le locus 4q35	42
Figure 20 : Structure d'une répétition <i>D4Z4</i>	46
Figure 21 : Représentation de l'ORF <i>DUX4</i> dans les répétitions <i>D4Z4</i> sur le chromosome 4q et 10q	47
Figure 22 : Schéma récapitulant les isoformes de <i>DUX4</i> produites selon les cellules	49
Figure 23 : Récapitulatif des fréquences de détection de <i>DUX4</i> publié	50
Figure 24 : Modèle d'expression pour <i>DUX4</i>	51
Figure 25 : Expression de <i>DUX4</i> et des marqueurs myogéniques au cours de la différenciation	53
Figure 26 : Modélisation de l'effet de compétition entre <i>DUX4-FL</i> et <i>DUX4-S</i> pour leurs cibles	55
Figure 27 : Localisation et transcription des miARN intergéniques et intragéniques	85
Figure 28 : Maturation des pri-miARN en pré-miARN par le complexe micropocesseur	86
Figure 29 : Un modèle du processing des micro ARN introniques	89
Figure 30 : Schéma récapitulatif des voies de biogénèse des miARN canoniques et non canoniques	90
Figure 31 : Position du miARN prédit sur <i>DUX4</i> et structuration de l'intron	91
Figure 32 : Récapitulatif des sondes utilisées pour détecter le miARN intronique par Northern Blot	92
Figure 33 : Northern Blot sur les cellules fœtales contrôle et FSHD	93
Figure 34 : Stratégie visant à déterminer la classe du miARN détecté par utilisation de ShARN ciblant les voies clés de maturation	95
Figure 35 : Stratégie visant à déterminer si le miARN est fonctionnel	96
Figure 36 : Stratégie visant à déterminer si le miARN peut interagir avec la 3'UTR de <i>DUX4</i>	96
Figure 37 : Récapitulatif des formes de <i>DUX4</i> exprimé et de l'état de méthylation H3K9 dans les différents lignages	101
Figure 38 : Schéma de la dernière répétition <i>D4Z4</i> sur le chromosome 4 et des sites d'épissages alternatifs	102

Figure 39 : Start alternatif et codon initiateur dans l'ORF de <i>DUX4</i> pouvant générer plusieurs isoformes de <i>DUX4</i> visible par sur-expression en western blot	105
Figure 40 : Modèle hypothétique d'expression de <i>DUX4</i> en lien avec les vagues myogéniques lors du développement	109
Figure 41 : Quantification du nombre de noyaux MyoD positifs par immunofluorescence sur des cellules fœtales FSHD et contrôle au cinquième jour de différenciation	110
Figure 42 : Proposition du modèle d'expression de <i>DUX4</i> et de la propagation du phénotype moléculaire dans un muscle entier	115
Figure 43 : Quantification du niveau d'expression de <i>FAT1</i> par RT-qPCR dans les cellules fœtales FSHD exprimant ou non un ShARN dirigé contre <i>DUX4</i> en différenciation	117
Figure 44 : Voies d'activations du programme myogénique en fonction du type musculaire	119

INTRODUCTION

1. Historique

La dystrophie Facio-Scapulo-Humérale (FSHD) est une pathologie à transmission autosomale à caractère dominant. Elle est la troisième pathologie musculaire la plus fréquente après la myopathie de Duchenne et la dystrophie Myotonique. Sa prévalence est de l'ordre de 1 naissance pour 20 000 (Orphanet). Néanmoins, compte tenu de l'hétérogénéité d'apparition des symptômes, elle est probablement sous-évaluée.

Cette pathologie fut caractérisée en 1885 par 2 médecins Français, Louis Landouzy et Jules Dejerine, qui ont identifié une population d'individus présentant les principales caractéristiques de la FSHD : une atteinte asymétrique et une faiblesse musculaire notamment au niveau du visage et des épaules (Jacqueline, 2011). Ce n'est qu'en 1950 que le critère génétique fut confirmé par Tyler en étudiant 17 familles dont la plus grande, constituée de 150 membres, est atteinte depuis 6 générations (Tyler, 1950).

2. Physiopathologie de la dystrophie Facio-Scapulo-Humérale

2.1 Une atteinte du muscle squelettique

En fonction du rôle et de son fonctionnement, le muscle peut être classé en deux groupes. Le muscle lisse qui tapisse la majorité des organes creux (comme le tube digestif) est relié au système vagal et se contracte donc de manière involontaire. Il permet ainsi la circulation des substances dans l'organisme. Le muscle strié se subdivise en muscle squelettique et muscle cardiaque.

Le muscle squelettique se compose de nombreuses myofibrilles, structures tubulaires allongées d'un diamètre de 1 à 2 μm , constituées de myofilaments disposés selon une organisation géométrique en trois dimensions extrêmement rigoureuse. Relié au système nerveux central par la jonction neuro-musculaire, il se contracte de manière volontaire (contrairement au muscle cardiaque qui répond à un stimulus autonome).

La formation du muscle squelettique passe par une série d'étapes qui démarre par l'engagement de précurseurs totipotents d'origine mésodermique vers le lignage musculaire. Ces précurseurs musculaires vont proliférer puis l'activation du programme myogénique permettra leur fusion en myotubes et la maturation de ces derniers pour former les fibres musculaires. Une fois que les fibres musculaires sont formées, un groupe de cellules

localisées à la périphérie de la fibre, les cellules satellites, assure les fonctions de régénération musculaire en cas de lésions. Elles apparaissent en fin de myogenèse fœtale et se positionnent sur la fibre sous la membrane basale. Au repos, les cellules satellites adultes expriment Pax7 (facteur de transcription) et sont dans un état quiescent, c'est-à-dire en arrêt mitotique (Kuang et al., 2008; Lagha et al., 2008; Buckingham and Rigby, 2014).

Lors d'un stress musculaire, tel qu'un effort physique important ou une lésion musculaire, la cellule satellite va passer du stade quiescent à un stade activé en perdant progressivement l'expression de PAX7 et en activant l'expression de Myf5 et MyoD (marqueurs myogéniques) (Kuang et al., 2008; Lagha et al., 2008; Buckingham and Rigby, 2014). Elle va alors réaliser une division asymétrique, une des cellules filles va s'engager vers la différenciation en exprimant MyoD et la deuxième cellule va ré-exprimer Pax7+ afin de retourner dans un état quiescent pour assurer l'autorenouvellement des cellules satellites au sein de la niche (Collins and Partridge, 2005; Kuang et al., 2007) (**Figure 1**).

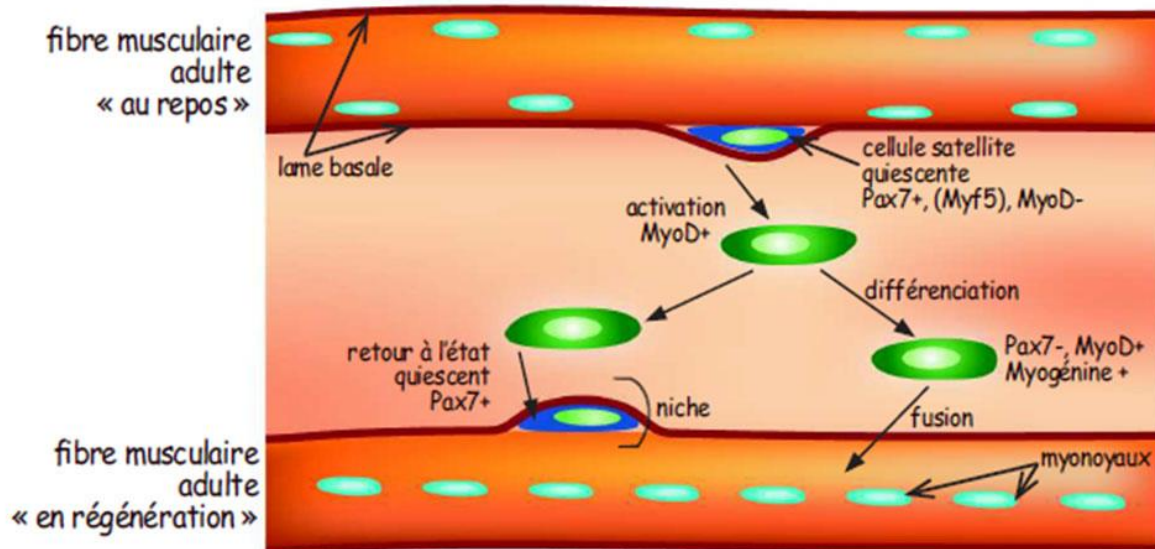


Figure 1 : Régénération du muscle squelettique (Jory et al., 2007)

2.2 Atteintes musculaires des patients FSHD

La FSHD est une pathologie progressive caractérisée par une fonte asymétrique des muscles du visage, des épaules et des bras induisant une faiblesse musculaire (**Figure 2**).

Les muscles du visage classiquement atteints sont l'obicularis oculi et l'obicularis oris. Le mouvement des paupières ainsi que des lèvres en est rendu difficile forçant les patients à dormir les yeux partiellement ouverts et figeant une partie de leurs expressions faciales (sourire). Bien que cette caractéristique soit l'un des critères de diagnostic, elle n'est retrouvée que chez 50% des patients.

Les muscles de la zone scapulaire, comme le trapèze ou le rhomboïde, font partie des muscles précocement atteints et sont donc les principaux symptômes conduisant à une consultation. Leurs atteintes se traduisent par une difficulté à lever les bras à l'horizontale au-dessus des épaules ainsi qu'à une saillie des omoplates avec attitude d'épaule tombante et en avant (décollement scapulaire).

Les muscles proximaux comme le biceps et le triceps sont atteints provoquant une atrophie du bras d'aspect en « bras de Popeye » (Harold, 2006). De plus, une atteinte sélective des muscles abdominaux entraînant une protubérance de l'abdomen vers l'avant, par remontée de l'ombilic (signe de Beever), est observée chez 90% des patients (Awerbuch and Wishnow, 1990). Les muscles respiratoires ne sont pas touchés par la pathologie. L'atteinte des muscles des membres inférieurs (tibialis antérieur) est progressive et provoque des difficultés à relever le pied ou à marcher. Dans environ 10 à 15% des cas, à un stade évolué, marcher ou se lever devient impossible et les patients sont contraints d'utiliser un fauteuil roulant.

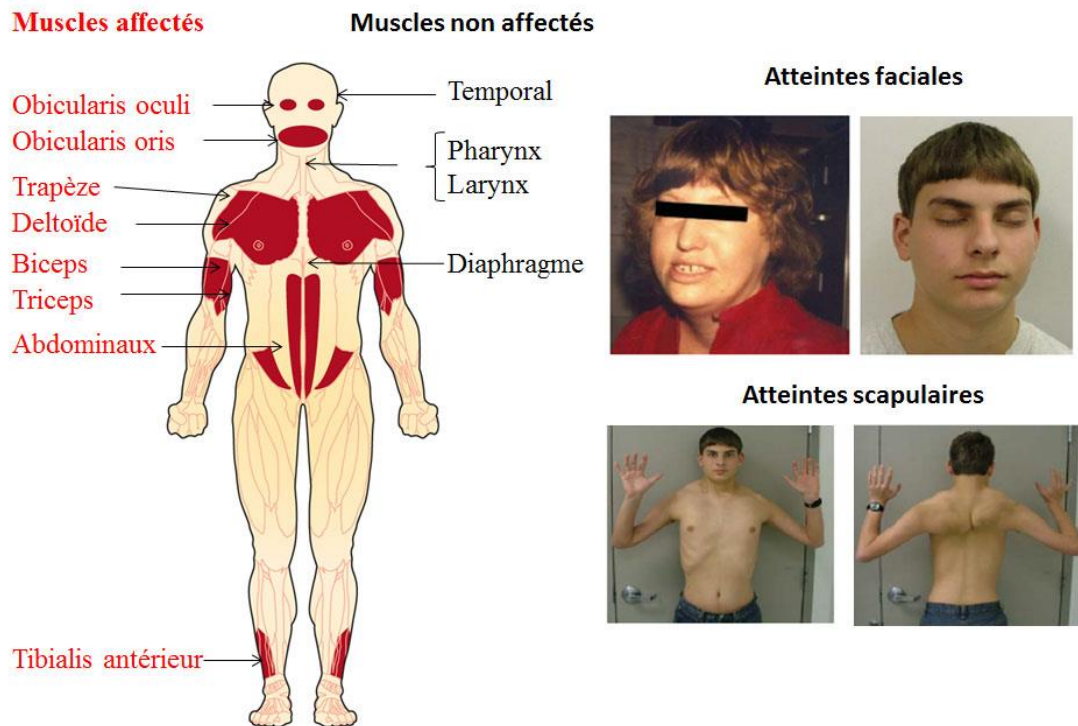


Figure 2 : Planche anatomique et photos de patients récapitulant la répartition des principaux muscles touchés ainsi que les signes caractéristiques de la dystrophie Facio-Scapulo-Humérale.

La FSHD est une pathologie évolutive apparaissant généralement au cours de l'adolescence avec une pénétrance de 95% à l'âge de 20 ans. Il semblerait que la pathologie soit plus fréquente chez les hommes que chez les femmes (Lunt et al., 1989; Tonini et al., 2004). Les premiers symptômes sont souvent ignorés (faiblesse du visage) et ce n'est que lors de l'atteinte de la zone scapulaire (handicapante) que les patients sont diagnostiqués. Généralement, les formes infantiles sont synonymes de sévérité. Bien que les critères phénotypiques de la FSHD soient identifiés, la forte hétérogénéité d'atteinte entre les patients la rend difficile à identifier. En effet, il a été décrit des cas de jumeaux homozygotes, possédant le génotype FSHD, où l'un des deux a développé les symptômes musculaires et est en fauteuil roulant tandis que l'autre est asymptomatique (Sakellariou et al., 2012).

2.3 Atteintes non musculaires

Un certain nombre de symptômes sans relation avec le muscle squelettique ont été décrits. Les principaux étant une diminution de l'audition, pour certaines fréquences, dans 64% des cas (Brouwer et al., 1991) ainsi qu'une vasculopathie dans 49 à 75% (Padberg et al., 1995). Les vasculopathies sont principalement asymptomatiques mais dans de très rares cas elles peuvent aboutir au développement d'une télangiectasie maculaire (maladie de Coat) correspondant à une augmentation du nombre de vaisseaux sanguins à l'intérieur de l'œil. Dans les formes les plus graves, cela aboutit à un détachement de la rétine et un endommagement définitif de l'œil.

Des défauts de la conduction cardiaque, avec une prédilection pour les arythmies auriculaires, ont également été décrits chez 5% des sujets (Faustmann et al., 1996). Une hypoventilation nocturne est également retrouvée dans 1% des cas, or les muscles respiratoires ne sont pas atteints dans la FSHD. Il semblerait que ce phénomène soit plus fréquemment retrouvé chez les patients qui ont développé un signe de Beavor (Wohlgemuth et al., 2004).

D'autres signes non spécifiques à la FSHD, mais associés à une atteinte du muscle squelettique, sont également utilisés dans le diagnostic clinique. En effet, il a été mis en évidence une variation du niveau de créatine kinase circulant pouvant aller jusqu'à 5 fois le taux habituel (0 à 195 UI/L) (Lunt and Harper, 1991), et un électromyogramme (EMG) caractéristique d'une myopathie (Stubgen, 2007).

Enfin une grande partie des patients font état de douleurs récurrentes ainsi que d'une fatigue accrue (Kalkman et al., 2006, 2008)

3. Aspect génomique de la FSHD

3.1 Identification d'un motif subtélomérique

Le locus de cette maladie a été localisé pour la première fois en 1990 (Wijmenga et al., 1990) puis sa position a été affinée en 1992 (Upadhyaya et al., 1992). Il se situe sur le bras long du chromosome 4 en position 4q35 à proximité de la région télomérique (Wijmenga et al., 1991). Plus tard, l'analyse de cette région a permis d'identifier un motif micro satellitaire

répété de 3,3 Kb , appelé *D4Z4*, orienté vers les télomères et dont la taille est moindre chez un groupe de patients appelé FSHD 1 (van Deutekom et al., 1993).

Ainsi, chez les individus sains, le nombre de répétitions *D4Z4* est compris entre 11 et 150 (Hewitt et al., 1994) tandis que les patients FSHD 1 présentent un nombre de copies compris entre 1 et 10 (**Figure 3**).

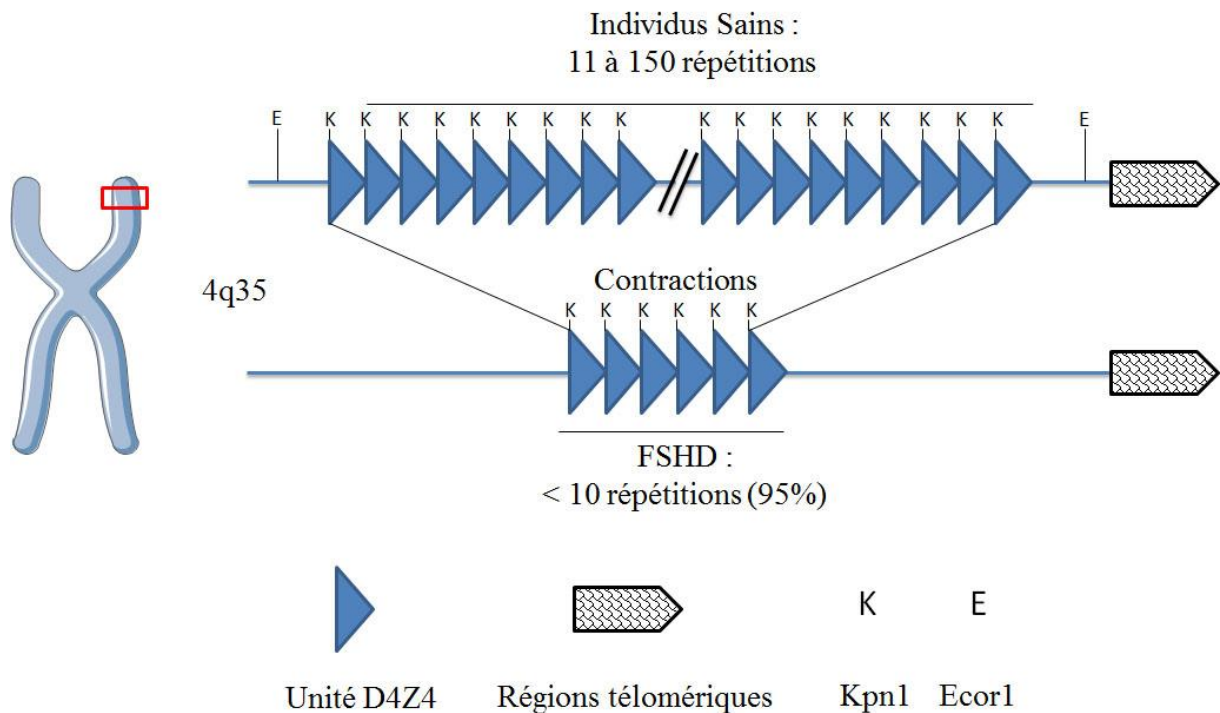


Figure 3 : Motif *D4Z4* sur le bras long du chromosome 4 en 4q35 chez les individus sains et atteints de FSHD : Chaque sous unité *D4Z4* de 3,3 kb est bornée par le site de restriction Kpn1, 95% des patients FSHD possèdent une contraction du motif répété *D4Z4*.

Il semblerait qu'il y ait une corrélation inverse entre le nombre de répétitions et la sévérité de la maladie. En effet, des patients présentant de grandes contractions *D4Z4* (1 à 3 répétitions) développent les cas les plus sévères avec une apparition durant l'enfance. Plus fréquemment, les patients possédant 4 à 7 répétitions *D4Z4* sont ceux montrant une plus grande variabilité inter-individus tandis que les patients ayant de 8 à 11 répétitions développent une FSHD modérée pouvant passer inaperçue (Lunt et al., 1995; Zatz et al., 1995; Salort-Campana et al., 2012).

La présence d'au moins une répétition semble indispensable au développement de la FSHD. En effet, un patient possédant une délétion totale de la région 4q35 (absence de répétition *D4Z4*) n'a montré aucun signe phénotypique caractéristique de la FSHD (Tupler et al., 1996).

Enfin, la diminution du nombre de ces répétitions n'est pas retrouvée chez tous les patients. En effet, environ 5% des individus développent une FSHD cliniquement identique, appelée FSHD 2, chez lesquels le nombre de répétitions *D4Z4* est normal suggérant alors l'implication d'autres facteurs génétiques.

3.2 Le motif *D4Z4* est également présent sur le chromosome 10

La région subtélomérique contenant les répétitions *D4Z4* est soumise à une forte instabilité. En effet, les répétitions *D4Z4* sont homologues à 98% avec celles du chromosome 10 (Bakker et al., 1995; Deidda et al., 1995) pouvant expliquer les nombreuses translocations entre ces 2 chromosomes dans la population générale (van Overveld et al., 2000). L'utilisation de sites de restriction spécifiques des répétitions du chromosome 4 et 10 ont permis de mettre en évidence la présence de translocation partielle aboutissant en une région subtélomérique hybride contenant à la fois des répétitions *D4Z4* du chromosome 10 et des répétitions issues du chromosome 4 (van Deutekom et al., 1996; van Overveld et al., 2000).

Malgré cette forte homologie, les contractions issues du chromosome 10 ne sont pas pathogéniques même dans les cas où les répétitions proviennent du chromosome 4 suggérant donc que les contractions ne sont pathogènes que sur un fond génétique lié au chromosome 4 (Cacurri et al., 1998; Lemmers et al., 1998).

3.3 Patients mosaïques

Le mosaïcisme cellulaire est caractérisé par la coexistence au sein d'un même tissu de plusieurs populations de cellules, souvent 2, possédant un génotype différent (notamment pour une mutation). Le degré de mosaïcisme s'explique par l'apparition de la mutation au cours du développement : plus la mutation survient tôt dans le développement de l'embryon, plus la proportion de cellules mutées est importante.

Selon le moment de survenue de la mutation, on parlera de mosaïque somatique (origine au début de l'embryogenèse) ou de mosaïque germinale (mutation apparue dans les gamètes).

L'exemple de mosaïcisme le plus répandu est l'inactivation de l'un des deux chromosomes X d'une fille conduisant à un mosaïcisme cellulaire si l'individu est hétérozygote pour une mutation du chromosome X (Migeon, 2006) car on observera deux

types de cellules selon le chromosome X qui a été inactivé. Le mosaïcisme fonctionnel créé par l'inactivation de l'X chez une femme hétérozygote peut être vu lorsque la mutation a un phénotype particulier. Dans le cas de l'albinisme oculaire, on observera au fond d'œil la présence de groupes de cellules pigmentées et dépigmentées selon que le chromosome X inactivé est ou non porteur de la mutation.

Dans la FSHD, la forte homologie des répétitions *D4Z4* avec le chromosome 10 ainsi que la localisation subtélomérique serait à l'origine d'une forte instabilité mitotique. Ainsi presque la moitié des cas de novo (40%) proviennent d'un réarrangement de la région *D4Z4* pendant la mitose, entraînant un mosaïcisme (van der Maarel et al., 2000). C'est probablement la raison pour laquelle parmi une cohorte de 208 individus sains issus de la population hollandaise, environ 3% présentent des réarrangements (contraction ou expansion) des répétitions *D4Z4* sur les chromosomes 4q ou 10q (van Overveld et al., 2000) contribuant à la variabilité du nombre de répétitions *D4Z4* dans la population générale.

L'utilisation de cellules issues de patients mosaïques offre un très bon modèle pour l'étude des dérégulations géniques chez les patients FSHD. En effet, le mosaïcisme au sein d'un tissu permet d'isoler à partir de la biopsie musculaire d'un patient FSHD, des cellules portant le génotype FSH (contractions *D4Z4* et chromosome permissif) et des cellules n'ayant pas les contractions (Krom et al., 2012). Brièvement, les cellules provenant d'une biopsie sont dans un premier temps isolées et purifiées selon leur potentiel myogénique puis immortalisées. L'immortalisation des myoblastes est réalisée en introduisant par vecteur viral le gène hTERT (télomérase), permettant l'élongation des télomères, et le gène CDK4 (Cyclin-Dependent Kinase) bloquant la voie de stress médié par P16^{INK4a} (Mamchaoui et al., 2011) (**Figure 4**). Le génotype des clones générés est ensuite vérifié par électrophorèse en champ pulsé (PFGE) (Lemmers et al., 2007).

Cette méthode permet ainsi de comparer deux populations de cellules ayant strictement le même génome hormis les contractions *D4Z4* sur le chromosome 4. De plus, puisque le degré d'atteinte est différent selon le muscle, ce modèle permet d'éliminer la variation inter-muscle (les clones proviennent de la même biopsie et donc du même muscle) et assure donc le meilleur contrôle interne possible.

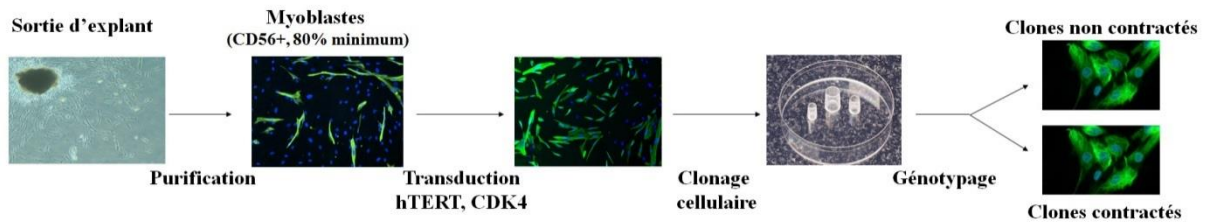


Figure 4 : Génération et immortalisation de clones cellulaires provenant de patients mosaïques : Les fragments de biopsies sont mis en culture et les cellules musculaires sortent de l'explant pour envahir progressivement le support. Un premier tri sur le marqueur membranaire cd56 est effectué afin de ne conserver que les myoblastes. Les cellules sont ensuite immortalisées par transduction de deux vecteurs viraux permettant l'expression de la télomérase (hTERT) et de CDK4. Chaque cellule est isolée (1 cellule par puits de culture) afin d'engendrer une population clonale génétiquement similaire sur laquelle le génotypage est effectué.

3.4 Diagnostic génétique par Southern blot

Puisque le diagnostic génétique est basé sur la mesure de la taille des répétitions *D4Z4*, dans un premier temps, l'ADN génomique est digéré par l'enzyme *EcorRI* suivi d'une détection de fragments *D4Z4* par southern Blot à l'aide d'une sonde (P13 e11) qui s'hybride avec la région immédiatement en amont de la première répétition. Ainsi, chez les individus contrôles, les fragments ont une taille comprise entre 40 et 350 kb tandis qu'elle est de 10 à 38kb chez les patients FSHD 1.

Puisque les répétitions *D4Z4* se retrouvent également sur le chromosome 10, une seconde étape de digestion enzymatique est nécessaire afin de les différencier. Ainsi l'utilisation des enzymes *XapI* (spécifique des répétitions issues du chromosome 4) et *BlnI* (spécifique des répétitions issues du chromosome 10) assure une bonne distinction. La taille des fragments étant relativement grande, l'utilisation d'une technique d'électrophorèse en champ pulsé (ECP) confère une meilleure résolution (Lemmers et al., 2004) (**Figure 5**).

Ce diagnostic n'est basé que sur la taille des répétitions *D4Z4*, or les patients FSHD 2 ne présentent pas de contraction rendant cette technique non applicable à ce groupe de patients. La recherche d'autres marqueurs de la FSHD a donc été nécessaire.

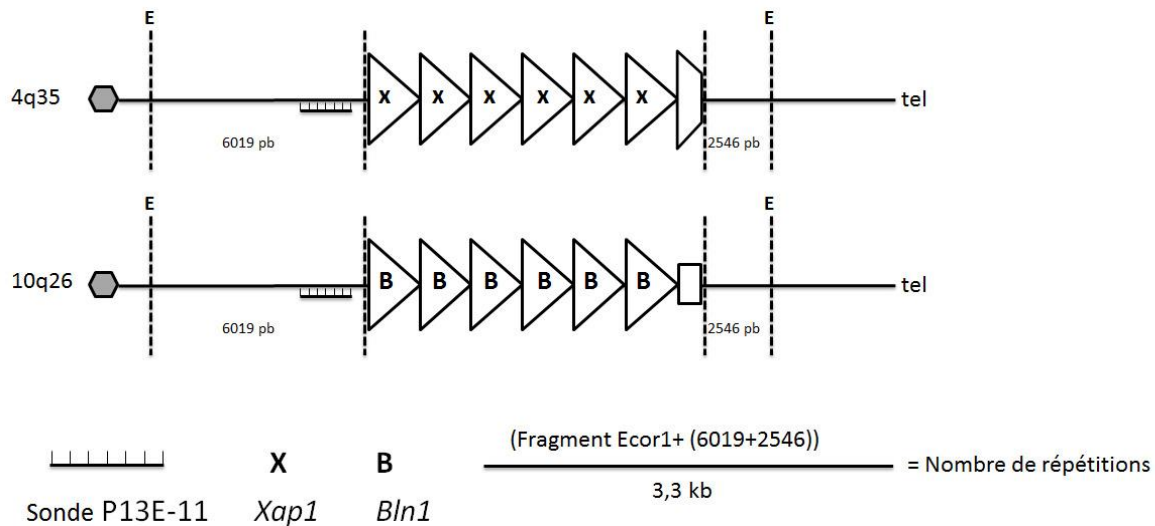


Figure 5 : Répartition des enzymes de restrictions sur les répétitions *D4Z4* entre le chromosome 4 et 10 ainsi que de la sonde P13E-11 utilisée pour mesurer la taille des répétitions par Southern Blot (adapté de Lemmers et al., 2001) : La digestion de l'ADN génomique par l'une des deux enzymes permet de faire la distinction entre les répétitions issues du chromosome 4 et du chromosome 10.

3.5 Complexité de la FSHD

L'existence des patients FSHD 2 (absence de contractions) cliniquement identiques au FSHD 1 ainsi que la nécessité de posséder au moins 1 répétition pour développer la pathologie ajoute un degré de complexité. En effet un patient possédant une délétion totale de la région 4q35 (absence de répétitions *D4Z4*) n'a montré aucun signe phénotypique caractéristique de la FSHD (Tupler et al., 1996). Prises ensemble, ces données suggèrent la présence d'éléments dans la région *D4Z4*, communs aux patients FSHD 1 et FSHD 2, nécessaire au développement de la pathologie.

3.5.1 Identification de différents segments polymorphes

Le séquençage de la région *D4Z4* a mis en évidence l'existence de segments polymorphes stables (SSLP : « simple sequence length polymorphism ») localisés à 3,5 Kb en amont de la première répétition (la plus proche du centromère) (**Figure 6**).

Plusieurs séquences polymorphiques ont été caractérisées, elles sont composées de 5 répétitions en tandem variable, d'une insertion/délétion de 8pb et de 2 SNP (« single

nucléotide polymorphism »). Certains de ces haplotypes sont ainsi préférentiellement retrouvés chez les patients FSHD comme l'haplotype 161 ou 166 (Lemmers et al., 2007, 2010; Scionti et al., 2012). Tout comme les répétitions *D4Z4*, la présence de ces haplotypes seuls n'est pas suffisante car il existe des patients porteurs de contractions *D4Z4* sur le chromosome 4 qui n'ont pas développé la pathologie (Galluzzi et al., 1999; Tonini et al., 2004). Ainsi, l'analyse d'une banque de YAC (« Yeast artificial chromosom ») par séquençage a mis en évidence l'existence de 2 autres haplotypes (exclusif) situés cette fois-ci immédiatement en aval de la dernière répétition (la plus proche des télomères) et appelés 4qA et 4qB (van Geel et al., 2002).

L'haplotype 4qA est constitué d'une séquence de 241 pb appelée pLAM contenant une séquence consensus ATTTAA associée à site de polyadénylation (PAS). Un polymorphisme au niveau d'un seul nucléotide (SNP, « Single Nucleotide Polymorphism ») est retrouvé dans cette séquence de polyadénylation dont certains directement associés à la pathologie (Lemmers et al., 2010). Cette région est immédiatement suivie de répétitions de type β -satellite de 68 pb sur une distance de 6,2 kb. L'allèle 4qB se termine quant à elle par les 570 premières pb d'une unité *D4Z4* suivi des séquences subtélomériques et télomériques (van Geel et al., 2002; Lemmers et al., 2002).

La notion de chromosome permissif a alors été introduite. Ainsi les patients FSHD 1 sont décrits génétiquement comme possédant un nombre de répétitions *D4Z4* compris entre 1 et 10, la présence de l'haplotype 4qA associé à l'un des SSLP décrit. Bien que les contractions *D4Z4* ne soient pas retrouvées chez les patients FSHD 2, la présence des différents haplotypes est nécessaire pour développer la FSHD (4qA et SSLP).

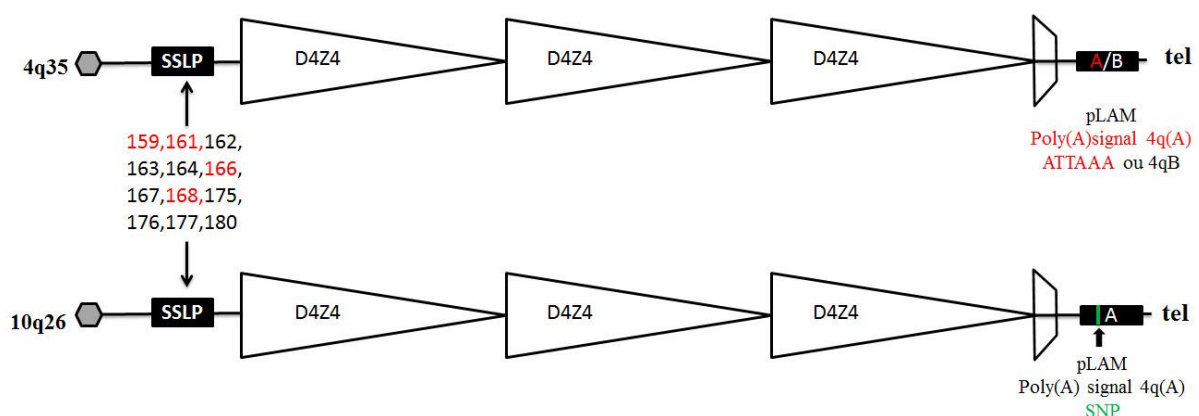


Figure 6 : Organisation des polymorphismes sur les répétitions *D4Z4* des chromosomes 4 et 10 (adapté de Scionti et al., 2012) : Les haplotypes permissifs associés à la FSHD sont représentés en rouge.

3.5.2 Prévalence des haplotypes

Bien que la combinaison de ces marqueurs génétiques soit indispensable au développement de la pathologie, la pénétrance de ces défauts génétiques est incomplète (**Figure 7**). En effet, l'analyse génétique sur un échantillon de 801 individus sains révèle que environ 3 % des sujets sains portent des allèles avec un nombre de répétitions *D4Z4* réduit (4 à 8) sur le chromosome 4q et que environ un tiers de ces allèles, 1,3 %, sont en association avec l'haplotype 4A161PAS (Scionti et al., 2012) et 0,8% sont associés à l'haplotype 4A166PAS. Enfin selon la même étude, sur 223 patients FSHD non liés, le génotype 4qA161-PAS est le plus représenté (53,7%) suivi du génotype 4qA166-PAS.

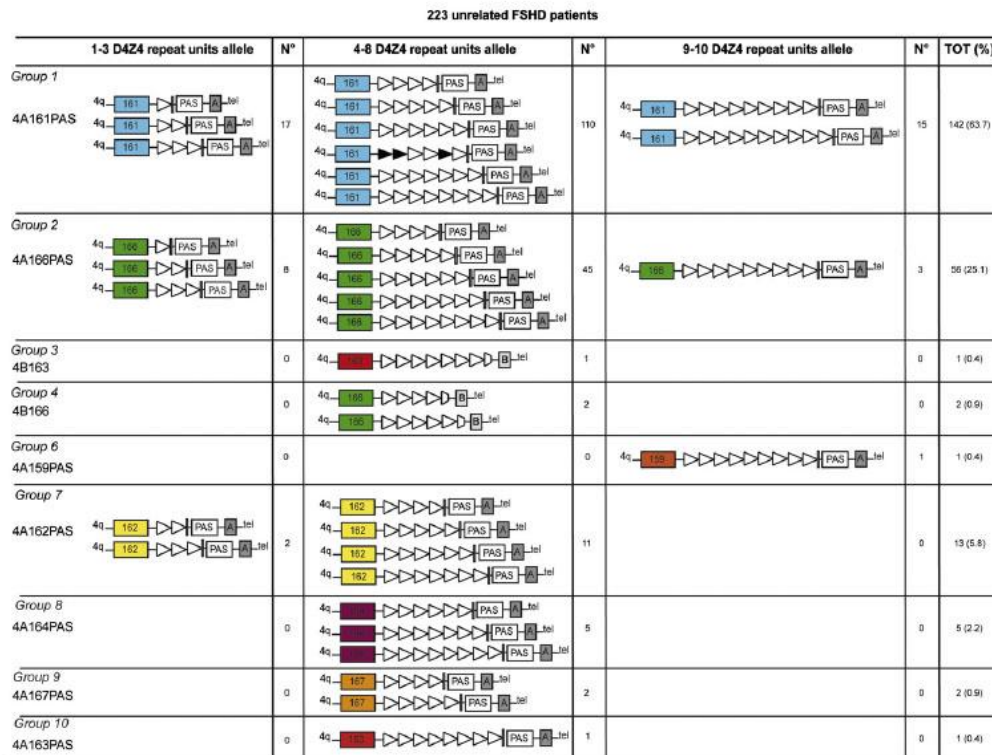


Figure 7 : Répartition des haplotypes au sein de 223 patients FSHD (Scionti et al., 2012) :

Représentation des haplotypes du chromosome 4 retrouvés chez 10% des 223 patients. Les haplotypes qui n'ont pas été observés chez ces patients sont : 4A166 (groupe 2), 4A162 (groupe 7), 4A164 (groupe 8), 4A167 (groupe 9) et 4A163 (groupe 10), 4B163 (groupe 3) et 4B166 (groupe 4). Les fréquences des haplotypes 4A161 (groupe 1, 63,7%) et 4A166 (groupe 2, 25,1%) sont différentes de celles précédemment rapportées. Les Chromosomes avec les haplotypes 4QB (groupes 3-4) ne possèdent ni pLAM ni PAS.

Bien que ces résultats suggèrent qu'il est possible d'avoir les caractéristiques génétiques de la FSHD sans développer de symptômes (2,1%), il n'est pas exclu que la pathologie ne se soit pas encore déclarée chez ces individus.

3.6 Diagnostic par peignage moléculaire

Environ 30% des patients FSHD sont des cas dits De Novo liés à des réarrangements d'origine mitotique (Bakker et al., 1996; van der Maarel et al., 2000) provoquant dans 40% de cas un mosaïcisme cellulaire pour les répétitions *D4Z4*. La présence de ce mosaïcisme ainsi que la nécessité de posséder certains haplotypes rend le diagnostic par southern blot difficile à interpréter.

Ainsi, afin d'améliorer la résolution du génotypage, une nouvelle technique de diagnostic a été développée. Le peignage moléculaire est basé sur la cartographie de la région *D4Z4* à l'aide de sondes spécifiques pour chaque région. (Nguyen et al., 2011) (**Figure 8**).

Concrètement, une lame en verre traitée avec de la silice est plongée dans la solution d'ADN à analyser. Par une méthode mécanique, les molécules d'ADN sont progressivement étirées sur la lamelle afin de former des fibres alignées. Une fois l'ADN fixé, il est hybridé avec les différentes sondes permettant la cartographie (Nguyen et al., 2011). Ainsi, un premier jeu de 2 sondes, marquées en bleu, hybride les régions proximales des répétitions *D4Z4* communes entre le chromosome 4 et 10. Une sonde de 6 kb, marquée en rouge, s'hybride à la région télomérique. Une sonde rouge hybride les régions β -satellites spécifiques de l'haplotype 4qA et une sonde bleue est spécifique de la région 4qB. Enfin une sonde verte s'hybride sur toutes les répétitions *D4Z4*. Il est alors possible de distinguer la taille des répétitions, l'haplotype et de différencier les allèles 4q et 10q

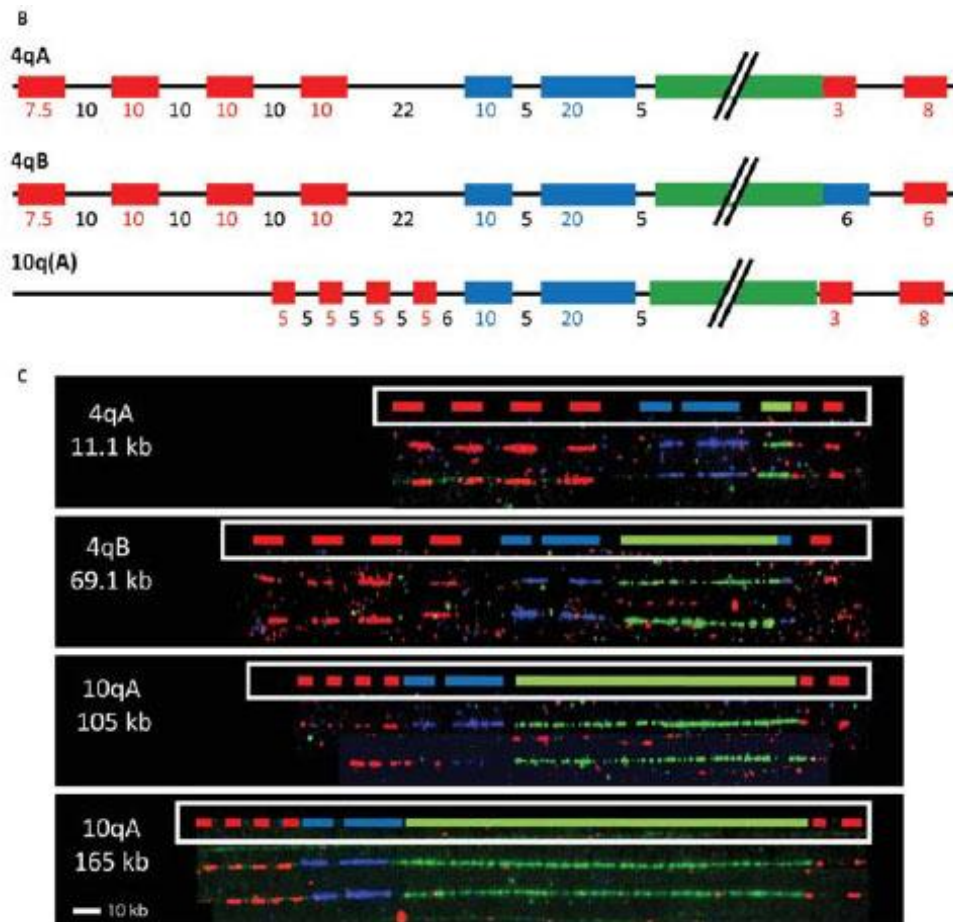


Figure 8 : Technique de peignage moléculaire (Nguyen et al., 2011b) :

B. Répartitions des sondes sur l'extrémité 4q35 et 10q26

C. Validation du code-barres 3 couleurs sur de l'ADN de patients. Un diagramme schématique de la visualisation attendue pour chaque allèle est présenté.

4. Modèles de régulations épigénétiques

4.1 Méthylation des régions *D4Z4*

Bien que les patients FSHD 2 ne présentent pas de contraction *D4Z4*, la nécessité de combiner certains des polymorphismes décrits est une caractéristique commune avec les patients FSHD 1. La recherche d'éléments communs entre les patients FSHD 1 et FSHD 2 a conduit à s'intéresser au statut épigénétique de la région *D4Z4*. L'analyse des répétitions *D4Z4* a montré la présence d'une forte proportion en GC semblable aux îlots CpG répartis dans le génome. L'étude de la méthylation de ces îlots a mis en évidence une

hypométhylation des répétitions *D4Z4* chez les patients FSHD1 comparés aux contrôles (van Overveld et al., 2003). Néanmoins, cette hypométhylation a également été retrouvée chez des individus contrôles portant des contractions *D4Z4* mais ayant un allèle non permissif (haplotype 4qB) laissant à penser que l'hypométhylation chez les patients FSHD 1 est associée aux contractions *D4Z4* (de Greef et al., 2009). Cette étude a également mis en évidence que, l'hypométhylation est plus importante chez des patients ayant un fragment *EcoR1* plus petit (10-20 kb) que dans un groupe de patients avec une contraction moyenne (20-31 kb), chez qui les symptômes étaient très variables.

Une autre étude a identifié la présence de différents marqueurs épigénétiques liés à un état d'hétérochromatine, caractérisé par la triméthylation des résidus H3K9 et H3K27, et des marques euchromatiniennes, caractérisées par la diméthylation des résidus H3K4 et l'acétylation de H3Ac (Zeng et al., 2009). La triméthylation des résidus H3K9 a été retrouvée fortement diminuée dans des cellules de patients FSHD 1 et FSHD 2. Les résidus H3K9me3 recrutent la protéine HP1 (Heterochromatin binding protein), qui joue un rôle important dans la répression transcriptionnelle (Bannister et al., 2001; Lachner et al., 2001), ainsi que des cohesines. La diminution des résidus H3K9me3 chez les patients FSHD induirait donc une perte des complexes répresseurs HP1-Cohesine et une déconformation de la chromatine (euchromatine) levant ainsi un mécanisme répressif des gènes localisés à proximité (**Figure 9**).

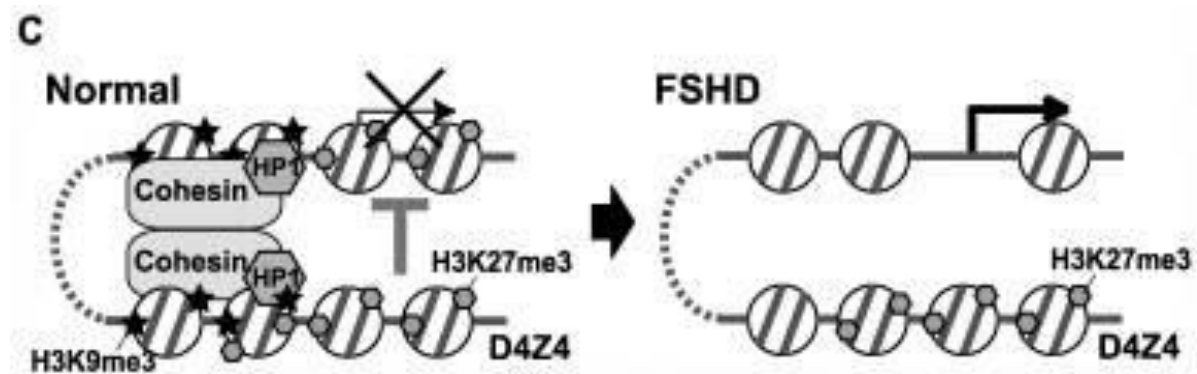


Figure 9 : Modèle schématique des changements chromatinien liés à la perte des H3K9me3 chez les patients FSHD (Zeng et al., 2009) : Perte des complexes répresseurs HP1-Cohesine et déconformation de la chromatine liée à la diminution des marqueurs H3K9me3.

4.2 L'hypométhylation est également retrouvée chez les patients FSHD 2

Sur le plan phénotypique, il n'existe aucune différence entre les patients FSHD 1 et FSHD 2 (van Overveld et al., 2003; de Greef et al., 2009, 2010). C'est pourquoi le profil de méthylation des répétitions *D4Z4* a été caractérisé chez ces patients. Ainsi tout comme les patients FSDH1, les répétitions *D4Z4* présentent une forte hypométhylation mais qui est plus marquée. Chez les patients FSHD 1, il semblerait que l'hypométhylation soit en lien avec la contraction du motif *D4Z4*, or, ce n'est pas le cas chez les patients FSHD 2 où elle est donc contraction indépendante (Lemmers et al., 2012). La cause de cette hypométhylation a donc été recherchée par séquençage d'exome et a ainsi permis d'identifier de rares mutations du gène *SMCHD1* localisées sur le chromosome 18 (« Structural Maintenance of Chromosomes flexible Hinge Domain-Containing Protein 1 » retrouvées chez 79% des patients FSHD 2 (Lemmers et al., 2012). Le panel des mutations est assez large et comprend de petites délétions, des mutations de sites d'épissages ou encore des mutations faux-sens aboutissant tous en une haplo insuffisance de la protéine (Lemmers et al., 2012). Ce gène est un modificateur chromatinien et permet l'établissement et le maintien de la méthylation des îlots CpG sur le chromosome X et certaines classes d'éléments répétés (Blewitt et al., 2008; Gendrel et al., 2012).

Des expériences d'immuno-précipitation de la chromatine ont permis de mettre en évidence une interaction physique entre *SMCHD1* et les répétitions *D4Z4* ainsi qu'une diminution de cette interaction chez les patients FSHD2 (comparés aux contrôles) (Lemmers et al., 2012). Tout comme les caractéristiques génétiques des patients FSHD1, la pénétrance de cette mutation est incomplète, ainsi il a été mis en évidence que sur 26 individus ayant une hypométhylation des répétitions *D4Z4*, une mutation de *SMCHD1* et un haplotype permissif, 19% sont asymptomatiques. Cette proportion de porteurs cliniquement non affectés est similaire à celle observée chez les patients FSHD1 (van der Maarel et al., 2007) et confirme la pénétrance incomplète de cette pathologie (**Figure 10**).

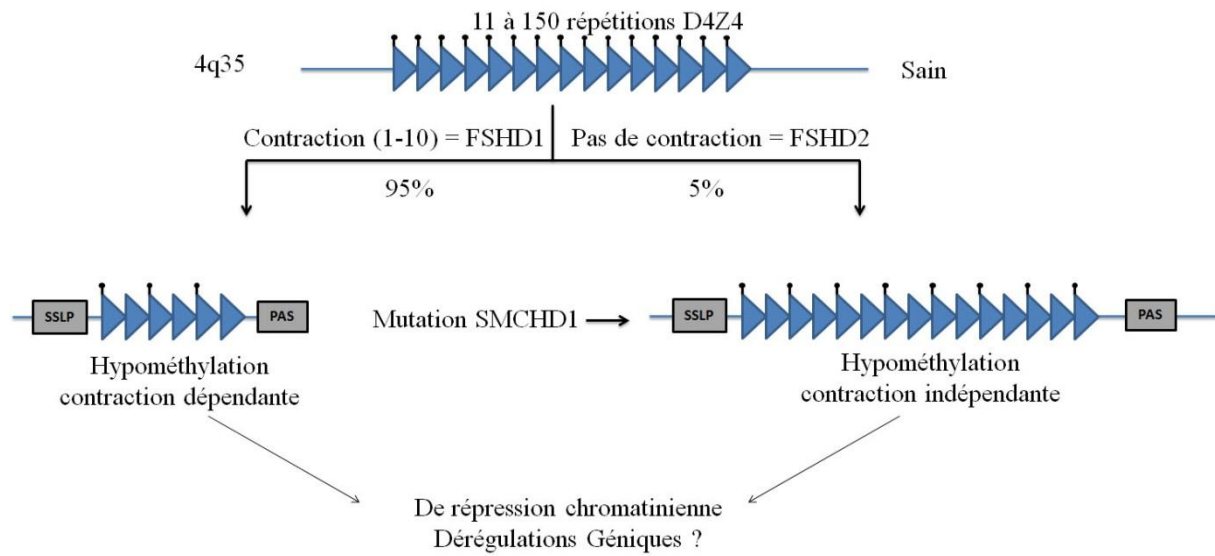


Figure 10 : Récapitulation des différences entre les patients FSHD 1 et FSHD 2.

4.3 Implication des télomères (TPE) et effets de positions

Les télomères sont des séquences d'ADN non codantes répétées en tandem (TTAGGG) à l'extrémité des chromosomes. Très conservés, ils ont pour fonction la protection des chromosomes, et contribuent ainsi au maintien de l'intégrité du matériel génétique. A mesure des divisions des cellules somatiques, la longueur des télomères diminue progressivement jusqu'à atteindre un stade critique provoquant l'instabilité du chromosome et la senescence répliquative de la cellule. Dans les cellules germinales, l'action d'une enzyme (la télomérase) maintient la longueur télomérique, mais cette enzyme est inactivée dans les cellules somatiques. Compte tenu de la structure adoptée par les télomères ainsi que de la forte méthylation, la région est principalement condensée en hétérochromatine bloquant la transcription des gènes situés à proximité. Puisque les répétitions *D4Z4* sont localisées en position subtélomérique (25 à 60 kb), le rôle du raccourcissement télomérique (TPE) a été étudié dans la pathogenèse de la FSHD. Pour cela, des cellules primaires issues de patients FSHD et Contrôles ont été immortalisées par ré-expression de la télomérase (vectorisée dans un lentivirus) afin d'empêcher le raccourcissement télomérique et donc la sénescence répliquative. La cassette d'expression de la télomérase, stablement intégrée au génome de la cellule, peut cependant être excisée à volonté grâce à l'action de la recombinaison Cre, afin d'obtenir des populations ayant des tailles de télomères différentes.

Ainsi, il a été constaté que dans les cellules où les télomères sont courts, le niveau de transcription de plusieurs gènes comme *DUX4* ou *FRG2* est augmenté (Stadler et al., 2013). De plus, le niveau de transcription est inversement proportionnel à la longueur télomérique (Figure 11).

Il apparaît donc que la longueur télomérique influe directement sur l'expression de plusieurs gènes localisés en 4q35. Néanmoins, la taille des télomères entre des individus contrôles et FSHD n'est pas modifiée, suggérant un impact minime des télomères sur son expression In-vivo.

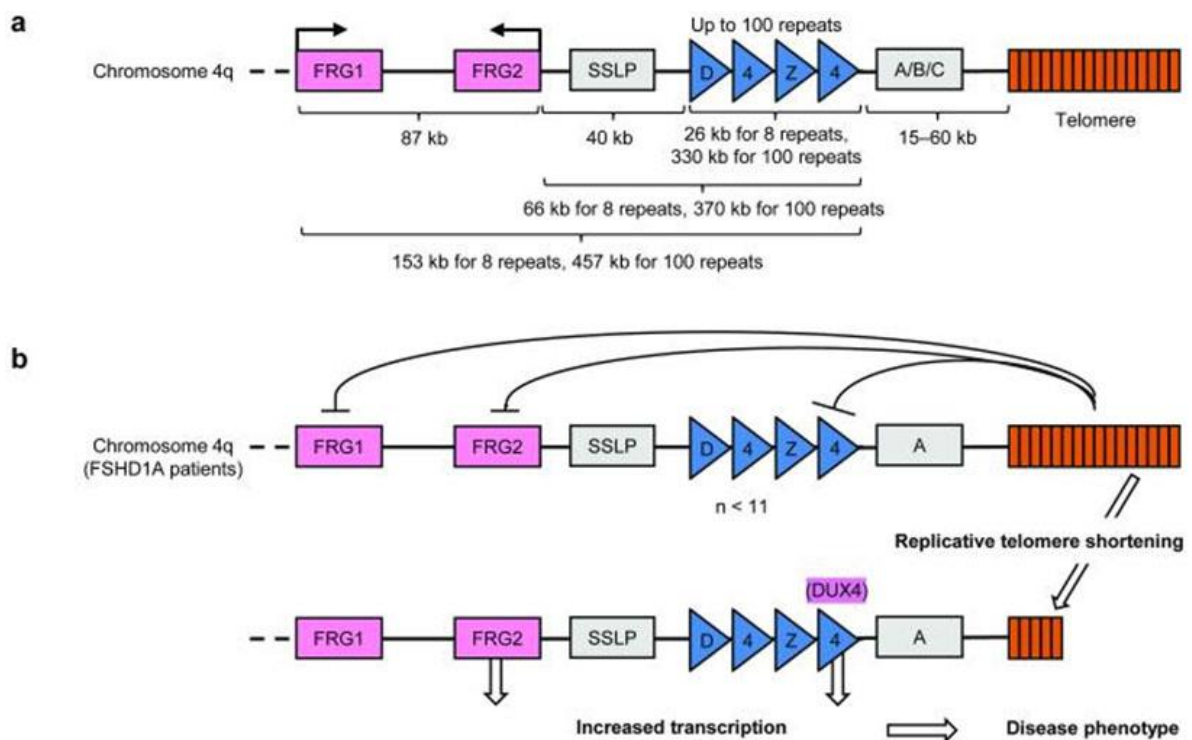


Figure 11 : Modèle de l'effet de position télomérique (TPE) sur la régulation des gènes du locus 4q35 (Stadler et al., 2013): La diminution du nombre de répétitions *D4Z4* associée à une diminution de la taille des télomères provoquerait une levée de la répression négative exercée par les télomères.

4.4 Domaine d'attache S/MAR et localisation nucléaire des répétitions D4Z4

Dans les noyaux des cellules eucaryotes ainsi que sur les chromosomes métaphasiques, l'ADN est organisé en domaine formant des structures en boucle (Buongiorno-Nardelli et al., 1982). Ces structures sont ancrées à la matrice nucléaire via des séquences spécifiques appelées domaine d'attache à la matrice nucléaire (S/MARs) (Mirkovitch et al., 1984; Cockerill and Garrard, 1986). Ces séquences sont classiquement retrouvées dans des régions non codantes (introns). De manières intéressantes, ce domaine d'attache est retrouvé à proximité des répétitions *D4Z4* (Petrov et al., 2006). Par des techniques d'immuno fluorescence, il a été montré que les répétitions *D4Z4* sont capables de fixer la matrice nucléaire via leurs domaines S/MARs. Ainsi l'association de S/MAR avec la matrice nucléaire isole les répétitions *D4Z4* dans une boucle d'ADN distincte des gènes situés plus en amont sur le chromosome 4. Cette interaction a été montrée comme diminuée de 30% dans les cellules primaires FSHD, comparés aux contrôles, regroupant ainsi sur la même boucle d'ADN les répétitions *D4Z4* et les gènes à proximité pouvant conduire à une dérégulation génique des gènes situés à proximité. Les analyses de 3C (Capture de Conformation de Chromatine) permettent de déterminer la distance ainsi que les interactions entre 2 fragments d'ADN. Ces mesures sur des myoblastes FSHD ont mis en évidence que les contractions *D4Z4*, entraînant la perte du domaine d'attache nucléaire, favorisent les interactions entre la région contenant les haplotypes 4qA/4qB et plusieurs gènes du locus 4q35 situés jusqu'à 5Mb (Bodega et al., 2009; Pirozhkova et al., 2008) (**Figure 12**).

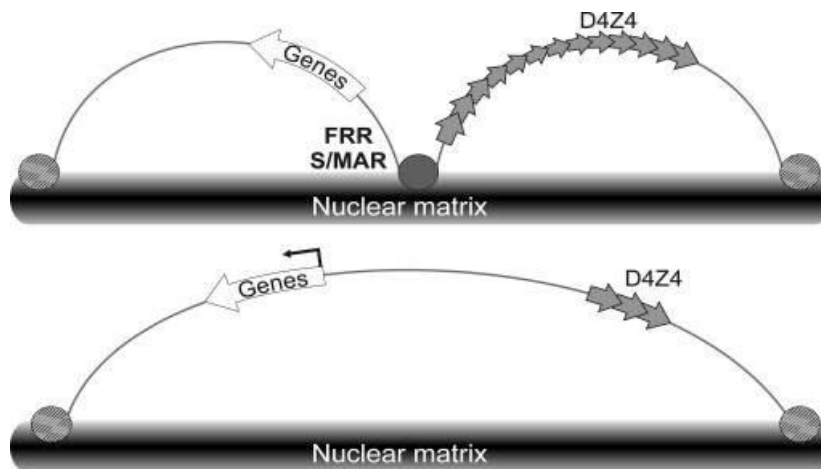


Figure 12 : Modèle d'attache à la membrane nucléaire et des boucles chromosomiques (Petrov et al., 2006)

La répartition de l'ADN dans les noyaux de cellules de mammifères n'est pas aléatoire. En effet, les noyaux sont compartimentés en domaine ayant des fonctions distinctes (transcription, épissage, réplication). Chaque chromosome occupe un emplacement précis qui semble être en relation avec la densité génique, l'activité transcriptionnelle et la réplication (Spector, 2001) (**Figure 13**). En effet, les régions d'euchromatine se retrouvent dans une zone interne du nucléoplasme (région riche en enzyme intervenant dans la transcription ou la réplication) tandis que les régions hétérochromatines sont plutôt retrouvées à la périphérie nucléaire (Shopland et al., 2003). Ainsi, l'architecture tri dimensionnelle du noyau semble jouer un rôle précis notamment au niveau de la régulation génique (Francastel et al., 2000; Cremer and Cremer, 2001; Jackson, 2003; Marshall, 2003). Il a été montré que l'extrémité 4q35 est préférentiellement localisée en périphérie nucléaire (85% dans les cellules musculaires) (Tam et al., 2004) et donc dans un domaine d'hétérochromatine. L'augmentation du nombre de répétitions D4Z4 semble légèrement altérer la localisation nucléaire des répétitions, ainsi, les grandes répétitions ont tendance à être moins périphériques (Ottaviani et al., 2010).

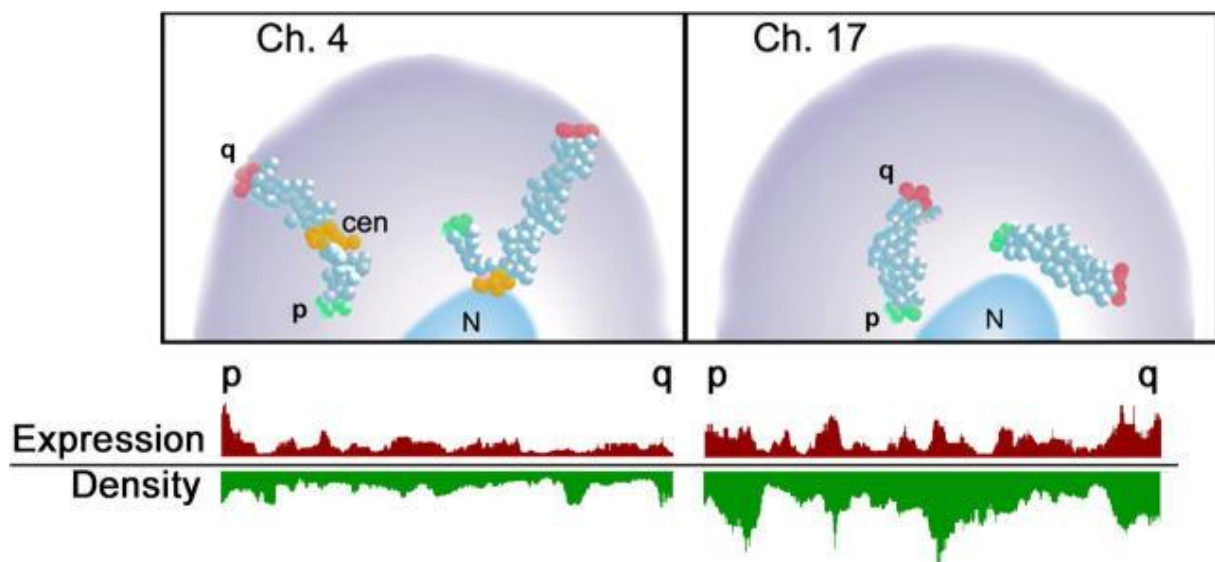


Figure 13 : Corrélation entre orientation spatiale et densité génique du chromosome 4 et 17 (Tam et al., 2004).

4.5 Identification de domaines présents dans *D4Z4*

La présence des répétitions *D4Z4* semble indispensable au développement de la FSHD. L'analyse du motif a mis en évidence l'existence de plusieurs domaines au sein des répétitions *D4Z4* qui peuvent participer à la régulation génique du locus 4q35 (**Figure 14**).

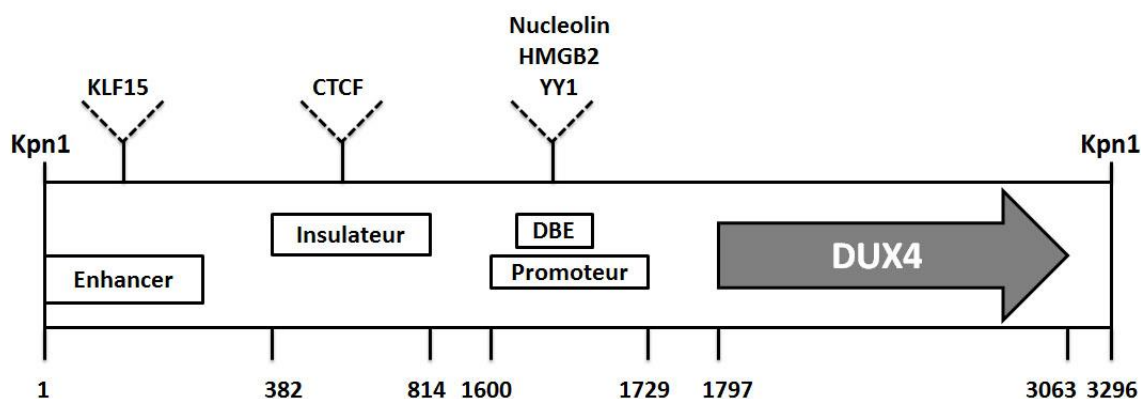


Figure 14 : Répartition des domaines présents dans les répétitions *D4Z4* (adapté de **Dmitriev et al.2011**) : Représentation schématique des éléments conservés à l'intérieur des répétitions *D4Z4*. Enhancer (ntd 1-329) (Petrov et al. 2007) contenant un site de fixation pour KLF15 (Dmitriev et al.2011) ; Insulateur (ntd 329-814) contenant un site CTCF (Ottaviani et al. 2009) ; Promoteur (ntd 1600-1729) contenant la TATA-box (CATAA) (Gabriels et al. 1999) ; Éléments de fixation *D4Z4* (DBE) contenant des sites pour la Nucleoline, HMGB2 et YY1 (Gabellini et al. 2002) ; Une phase ouverte de lecture pour DUX4 (ntd 1797-3063) (Gabriels et al. 1999).

4.5.1 Domaine enhancer

Les Enhancer sont de petites séquences ADN régulatrices capables de fixer des complexes protéiques afin de stimuler la transcription et cela à grande distance. La cartographie du motif *D4Z4* a ainsi mis en évidence l'existence d'un enhancer en position nucléotidique 1 à 329 (Petrov et al., 2003, 2008). Des constructions contenant l'enhancer associé à un gène rapporteur ont mis en évidence que l'activité transactivatrice était dépendante de la fixation de la protéine KLF15 (« KRUPPEL-LIKE FACTOR 15 ») (Dmitriev et al., 2011) (**Figure 15**). Il a été montré que KLF15 est surexprimé au cours de la différenciation musculaire et que l'activité enhancer est stimulée par le facteur myogénique MyoD. De plus, une surexpression de MyoD induit également l'expression de KLF15

suggérant donc que KLF15 associé au facteur myogénique MyoD permet une stimulation de l'activité enhancer des répétitions *D4Z4*. Le modèle proposé est ainsi le suivant, lorsque la structure chromatinienne est condensée (heterochromatine, absence de contractions) KLF15 ne peut pas se fixer à l'enhancer dans les répétitions et n'est donc pas activé. Lorsque le motif *D4Z4* se contracte, la perte des marqueurs hétérochromatines au profit de marqueurs euchromatines (hypométhylation H3K9) permet la fixation de KLF15 aux enhancer *D4Z4* et stimule alors leurs effets. Associé à la perte du motif d'attache à la membrane nucléaire MAR, le rapprochement dans la même boucle chromosomique des répétitions et des gènes du locus 4q35 pourrait induire leurs dérégulations médiées par l'activité de l'enhancer.

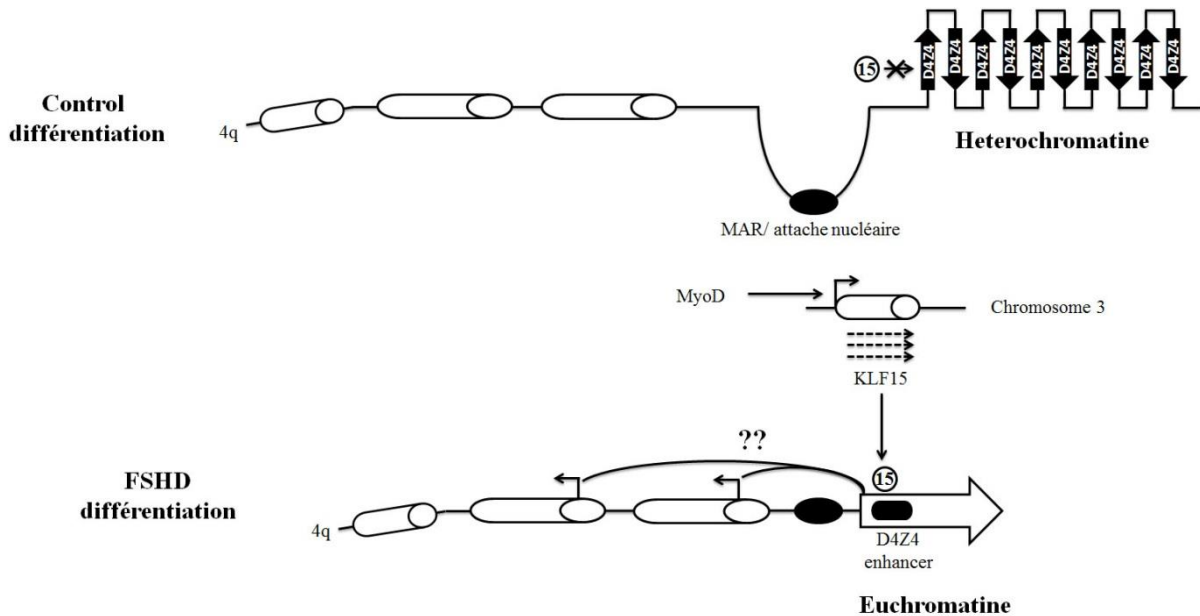


Figure 15 : Modèle schématique de l'activation de l'enhancer dans les répétitions *D4Z4* par le facteur KLF15 (adapté de Dmitriev et al.2011) : L'activation de MyoD au cours de la différenciation induit l'expression de KLF15. Chez les individus contrôles, la conformation fermée liée aux grands nombres de répétitions *D4Z4* bloque la fixation de KLF15 sur l'enhancer dans *D4Z4*. La relaxation de la chromatine localement facilite l'accès de KLF15 à l'enhancer permettant ainsi son activation et la stimulation de la transcription des gènes du locus 4q35

4.5.2 Domaine insulateur *D4Z4*

Les insulateurs sont des séquences d'ADN possédant deux propriétés distinctes. Ils protègent les régions transcriptionnellement actives des effets conformationnels de l'hétérochromatine, et interfèrent avec le promoteur et les séquences enhancer lorsqu'ils sont placés entre. Il apparaît ainsi que la présence en grand nombre des répétitions *D4Z4* aurait des propriétés de barrières chromatinienne (Ottaviani et al., 2009; Wijmenga et al., 1992).

L'étude des répétitions *D4Z4* a permis de conclure qu'au moins une répétition était capable de bloquer l'effet de la chromatine par effet de CPE (Chromatine Position Effect) et de bloquer la communication entre un promoteur et un enhancer (Ottaviani et al., 2009). Des expériences d'immuno précipitation de la chromatine (ChIP) ont mis en évidence la fixation de la protéine CTCF (protéine couramment impliquée dans les effets insulateurs) sur les répétitions *D4Z4*. De plus, dans des cellules contenant des constructions mimant l'organisation du locus *D4Z4*, l'extinction de CTCF par ARN interférence aboutit en une répression progressive de l'expression du gène rapporteur situé à proximité suggérant que la fixation de CTCF sur les répétitions a un effet anti-silencing. Une seconde interaction a ensuite été identifiée associant les répétitions *D4Z4*, CTCF et les lamines de types A/C permettant l'accrochage de la région contractée en périphérie nucléaire (Ottaviani et al., 2009) (**Figure 16**). Il semblerait que cette interaction n'ait pas lieu dans un contexte non contracté car la multimérisation des répétitions *D4Z4* diminue l'accessibilité de la région ainsi que le contexte d'hétérochromatine lié à une forte méthylation de la région. Dans cette situation, les répétitions *D4Z4* joueraient un rôle répresseur. Cependant, dans un contexte FSHD, la réduction du nombre de répétition *D4Z4* permet la liaison de CTCF et des lamines de type A à la périphérie nucléaire, provoquant alors un changement dans la fonction biologique de *D4Z4* qui passerait d'un répresseur à un insulateur.

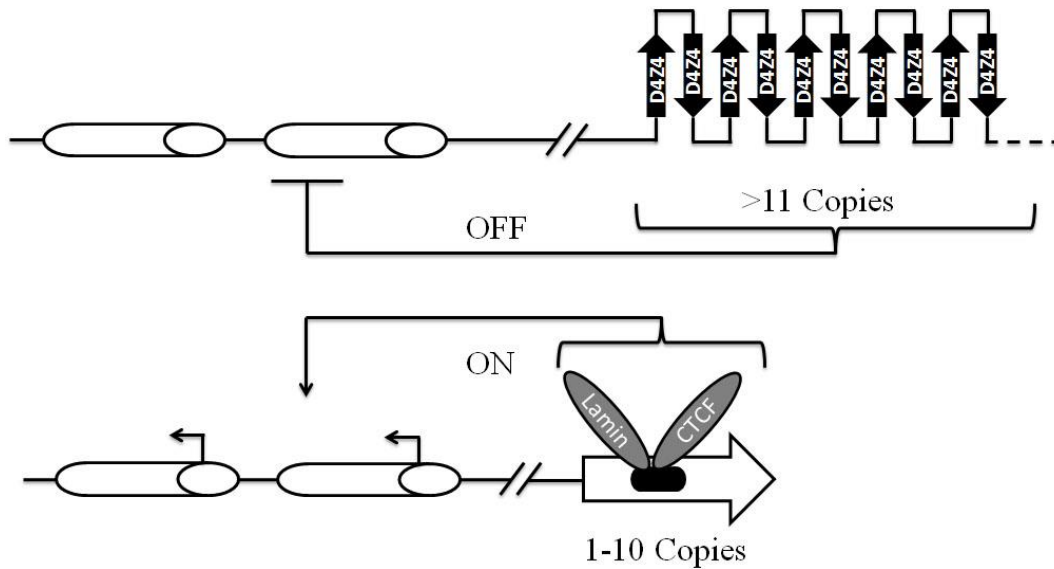


Figure 16 : Modèle expliquant le rôle de l'insulateur *D4Z4* et son implication dans les modifications épigénétiques de la FSHD (adapté de Ottaviani et al., 2009) : Chez les individus non contractés, la multimérisation des répétitions empêche la fixation de CTCF et des Lamines de type A/C. Dans cette conformation, les répétitions *D4Z4* ont une fonction répressive sur les gènes du locus 4q35. À l'inverse, chez les patients, la diminution du nombre de *D4Z4* permet la fixation de CTCF et des Lamines de type A/C transformant l'activité répressive des répétitions en activité insulatrice levant ainsi la répression des gènes du locus 4q35.

4.5.3 Domaine répresseur DBE

Couramment, la régulation de la transcription fait intervenir des complexes multi protéiques qui jouent un rôle dans la modulation de l'expression génique. Des expériences de retard sur gel ont permis de mettre en évidence que les répétitions *D4Z4* étaient capables de fixer un complexe multi protéique au niveau d'une séquence de 27 pb appelée DBE (*D4Z4* Binding Element) (Gabellini et al., 2002) (**Figure 17**). Ce complexe, jouant le rôle de répresseur transcriptionnel, est constitué de 3 protéines distinctes : un facteur de transcription appelé YY1 (Gordon et al., 2006), une protéine de liaison à l'ADN appelée HMGB2 (High Mobility Group Box 2) (Wanschura et al., 1996), et une protéine NCL (Nucleolin) impliquée dans le contrôle de la transcription (Srivastava et al., 1989). Ainsi, la fixation de ce complexe répresseur sur les répétitions *D4Z4* contribuerait à la répression transcriptionnelle de la région 4q35. Le modèle alors posé est le suivant : chez les patients FSHD1, les contractions *D4Z4*

ont pour conséquence une diminution du nombre de complexe répresseur sur la région et provoqueraient une levée de la répression des gènes situés à proximité.

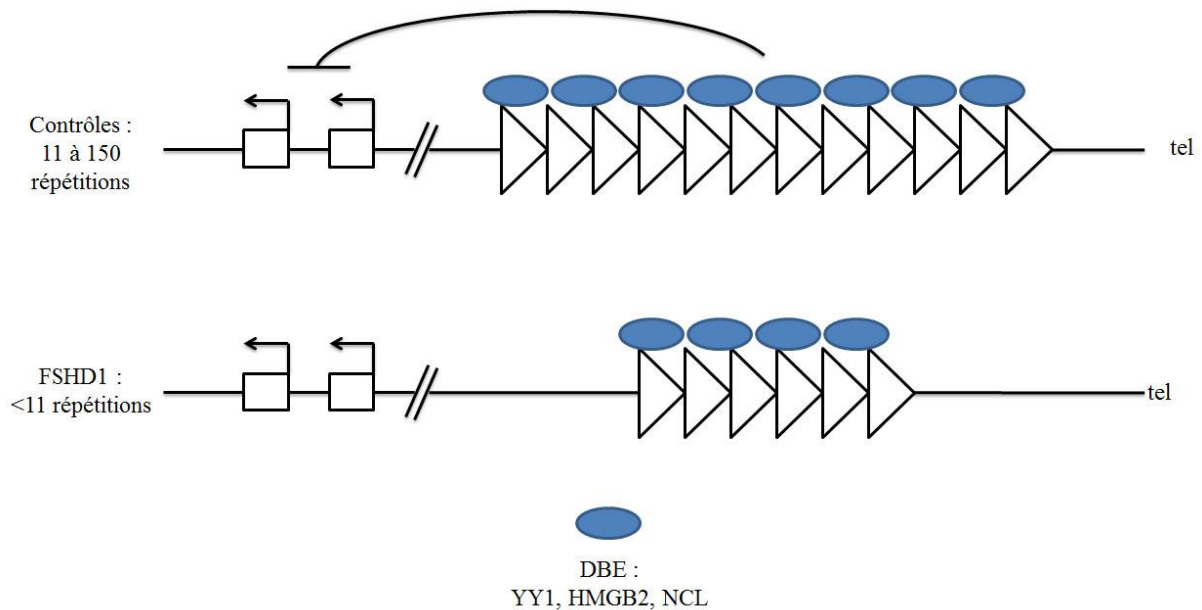


Figure 17 : Modèle de répression transcriptionnelle médiée par la fixation du complexe multiprotéique YY1, HMGB1 et Nucleoline sur le domaine DBE (adapté de Gabellini et al., 2002) : Les contractions *D4Z4* ont pour conséquence une diminution du nombre de complexes répresseurs fixés sur la région provoquant une levée de la répression transcriptionnelle des gènes à proximité.

Normalement, les répétitions *D4Z4* maintiennent un nombre suffisant de complexes répresseurs (Complexe polycombe) (Cabianca et al., 2012), assurant un blocage transcriptionnel de la région 4q35. Or les contractions *D4Z4* induiraient une dé-répression de la région, via la diminution de la quantité de répresseurs fixés, permettant ainsi la production d'un long ARN non codant (DBE-T, *D4Z4* Binding Element-transcrit). Cet ARN non codant serait produit à partir de la première répétition *D4Z4*, la plus proche du centromère, au niveau du domaine DBE (Cabianca et al., 2012). Cet ARN non codant, d'une taille de 2,8 kb, est capable de recruter un groupe de protéines de la famille des tritorax (protéines impliquées dans la prolifération cellulaire ou l'inactivation du chromosome X), (Schuettengruber et al., 2007), sur le locus FSHD (**Figure 18**). Ceci a pour conséquence un switch épigénétique en stimulant la diméthylation de H3K36, le remodelage chromatinien et la transcription de gènes sur le locus 4q35. De manière intéressante, ce transcrit est retrouvé spécifiquement sur-exprimé dans des muscles et cellules de patients FSHD suggérant donc qu'il est impliqué dans la pathologie.

Néanmoins, puisque les répétitions *D4Z4* sont retrouvées sur le chromosome 10, les auteurs de l'étude sur DBE-T ont travaillé sur des cellules ovariennes de hamster chinois(CHO) contenant soit le chromosome 4 humain (provenant d'un individu contrôle) soit transfecté avec un BAC (Bacterial Artificial Chromosome) possédant ou non les répétitions *D4Z4*. Ainsi, il est difficile de confirmer la production du transcrit DEB-T en l'absence d'une étude sur des cellules humaines car il apparaît de plus en plus que le contexte épigénétique d'une cellule joue un rôle majeur dans la FSHD.

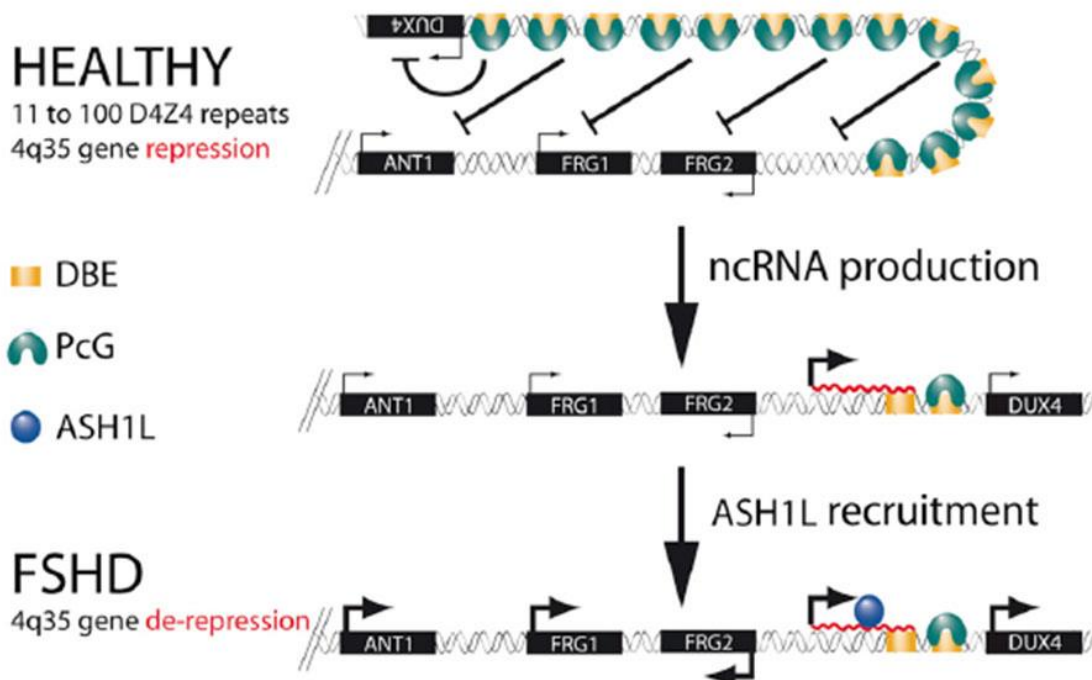


Figure 18 : Modèle de régulation épigénétique du locus 4q35 par les complexes polycombs/trithorax (Cabianca et al., 2012): Les individus sans contraction D4Z4 ont un grand nombre de protéines PcG associées aux répétitions assurant le maintien d'une chromatine répressive au locus 4q35. La contraction des répétitions a pour conséquence une diminution du nombre de PcG localement provoquant une expression de DBE-T. Ce long ARN non codant recrute alors la protéine ASH1L stimulant une réorganisation topologique du locus FSHD induisant alors une dé-répression des gènes à proximité.

4.6 Capacité bidirectionnelle du promoteur *D4Z4*

L'analyse d'éléments répétés dispersés dans le génome a mis en évidence que certaines régions présentaient une transcription sur le brin sens et anti-sens (Cho et al., 2005; Chadwick, 2008; Snider et al., 2009). L'analyse du motif *D4Z4* a mis en évidence l'existence d'un promoteur putatif TACAAA situé 149 pb en amont d'une phase ouverte de lecture (Hewitt et al., 1994; Lee et al., 1995; Winokur et al., 1994). Afin d'étudier l'activité de ce motif, la région promotrice a été insérée en amont d'un gène rapporteur ou entre deux gènes rapporteurs afin de mesurer l'activité transcriptionnelle sens et anti-sens (Block et al., 2012). Ainsi, il a été montré que la région promotrice est capable de conduire une activité de transcription sens et anti-sens simultanément. De manière intéressante, l'activité transcriptionnelle anti-sens est environ 4 à 5 fois plus importante que l'activité sens. De plus la transcription est stimulée lorsque les cellules sont mises en différenciation mais l'asymétrie de transcription reste inchangée. Enfin, les mêmes mesures ont été réalisées dans des cellules souches embryonnaires contrôles et montrent que l'activité sens est très fortement réprimée (1 cellule sur 250 exprime le transgène sens).

Il semble donc que l'état de la cellule (ES, myoblastes, myotubes) définisse le sens de la transcription et puisse jouer un rôle dans la régulation du locus *D4Z4* notamment par la production d'ARN anti-sens.

5. Gènes dérégulés dans la FSHD

Les nombreuses modifications épigénétiques du locus 4q35 communes aux patients FSHD 1 et FSHD 2, comme les changements de conformation chromatinienne ou encore les modifications de la méthylation (liées aux contractions ou à la mutation de SMCHD1), peuvent être à l'origine de dérégulations géniques en Cis (sur le même chromosome). Ainsi de nombreuses équipes ont regardé le niveau d'expression d'un certain nombre de gènes localisés à proximité des répétitions *D4Z4* (**Figure 19**).

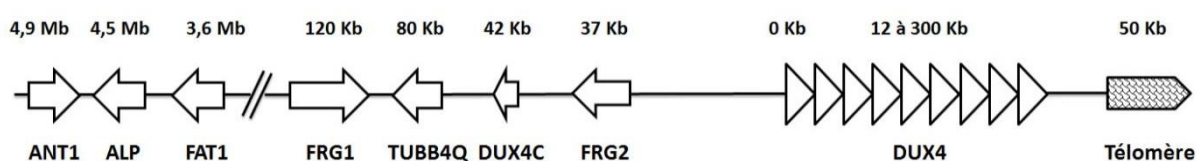


Figure 19 : Représentation des gènes étudiés dans la FSHD sur le locus 4q35.

5.1 FRG1

Ce gène fut le premier décrit comme dérégulé chez des patients FSHD (Gabellini et al., 2002). *FRG1* (FSHD Region Gene 1), localisé à 120 kb des répétitions, est constitué de 9 exons et produit une protéine de 258 amino acides intervenant dans la maturation des pre-messager (Kim et al., 2001; Rappsilber et al., 2002; van Koningsbruggen et al., 2004). Il a été montré que l'expression de *FRG1* est vitale pour le développement du muscle squelettique (Hanel et al., 2009) et l'angiogenèse (Wuebbles et al., 2009). En effet, il interagit avec l'Actine F, promeut sa polymérisation et permet ainsi la structuration des fibres musculaires (Liu et al., 2010). Néanmoins, la dérégulation de ce gène chez les patients FSHD est sujette à controverse. En effet, ce gène n'a pas été confirmé comme surexprimé, chez des patients adultes FSHD, par plusieurs laboratoires (Osborne et al., 2007; Masny et al., 2010), tandis qu'il a été retrouvé surexprimé dans des biopsies fœtales FSHD (Broucqsault et al., 2013).

Néanmoins, un modèle de souris sur-exprimant *FRG1* (Gabellini et al., 2006) induit un phénotype dystrophique. En effet, des analyses histologiques ont confirmé la présence chez cette souris de noyaux centraux (caractéristique d'un muscle lésé en régénération), une augmentation de la taille des fibres et une nécrose. Plusieurs études confèrent un rôle de

FRG1 dans la maturation des pré-messager (Kim et al., 2001; Rappsilber et al., 2002; van Koningsbruggen et al., 2004). Ainsi, l'analyse de l'épissage de certains transcrits décrit dans d'autres pathologies comme la dystrophie myotonique (Ranum and Day, 2004) a mis en évidence une altération du profil d'épissage de TNNT3 (Troponine T3).

L'étude approfondie de ce modèle a permis de comprendre la cascade de dérégulation aboutissant aux symptômes dystrophiques. Brièvement, *FRG1* est capable d'interagir avec l'ARN *Rbfox1* (RNA Binding protein FOX1) qui donne naissance à un facteur d'épissage (Fukumura et al., 2007) notamment retrouvé dans le cerveau, le muscle squelettique et le cœur (McKee et al., 2005; Kuroyanagi, 2009; Tang et al., 2009; Damianov and Black, 2010) et diminue alors la stabilité du transcrit. A son tour, *Rbfox1* impact sur l'épissage alternatif de certains gènes notamment l'exon 6 de la Calpaine 3 (impliqué dans la myogenèse). L'inclusion de l'exon 6 permet de générer une isoforme de la Calpaine 3 indispensable au développement et à la régénération musculaire (Beckmann and Spencer, 2008). Ainsi la souris sur-exprimant *FRG1* montre une diminution de *Rbfox1* qui provoque une augmentation de l'exclusion de l'exon 6 de la Calpaine 3 induisant des défauts au niveau musculaire.

Bien que cette souris soit présentée comme un « modèle de la FSHD » elle ne développe aucun autre symptôme spécifique de la FSHD comme l'atteinte asymétrique ou encore les défauts non-musculaires (vasculopathie). De plus, il est probable que la sur-expression d'une protéine spécifiquement dans le muscle squelettique induise un phénotype dystrophique sans lien avec la FSHD. Néanmoins, il apparaît de plus en plus que la pathologie n'est pas liée à une seule mutation mais qu'elle est probablement d'origine multigénique pouvant expliquer l'absence des signes cliniques spécifiques de la FSHD.

5.2 FRG2

Localisé à 37kb des répétitions *D4Z4* en 4q35, le rôle de *FRG2* (FSHD Region Gene 2) dans la FSHD a donc été étudié. Composé de 4 exons, et d'un promoteur supposé être muscle spécifique, produit une protéine de 278 aminoacides de fonction encore inconnue. Deux équipes ont montré que *FRG2* est progressivement sur-exprimé au cours de la différenciation de myoblastes issus de patients FSHD alors qu'il n'est pas détecté chez des individus contrôles ou atteints d'autres myopathies (Gabellini et al., 2002; Rijkers et al., 2004). Or, ce gène est également retrouvé sur le chromosome 10q et possède une forte homologie (4 nucléotides de différent sur les 3kb), l'origine du transcrit est donc incertaine.

Néanmoins, ces résultats sont très controversés car à ce jour aucune autre étude n'a confirmé cette dérégulation (Osborne et al., 2007; Masny et al., 2010). De plus, une souris sur-exprimant *FRG2* n'a développé aucun phénotype (Gabellini et al., 2006) suggérant tout comme *FRG1* que ce gène n'est pas impliqué dans la physiopathologie ou bien que sa dérégulation seule ne soit pas suffisante à l'établissement d'un phénotype FSHD.

Enfin il a été mis en évidence la présence chez plusieurs familles d'une délétion d'une grande partie de la région chromosomique 4q35 (Lemmers et al., 2003; Deak et al., 2007), ainsi, le rôle de *FRG2* dans la pathologie semble exclu.

5.3 ANT1

Localisé à grande distance des répétitions *D4Z4* (4,9 Mb) il code pour l'adénine nucléotide translocator 1 (*ANT1*) et est exprimé dans le cœur et le muscle squelettique. Il joue un rôle central dans l'export de l'ATP de la matrice mitochondriale et l'import de l'ADP. De par sa fonction, des mutations de ce gène provoquent un grand nombre de pathologies neuromusculaires comme l'Ophtalmoplégie ou encore la Cardiomyopathie hypertrophique (Sharer, 2005). Sa sur-expression est également très toxique (dans des cellules rénales embryonnaires Humaines), elle induit notamment un accroissement des ROS (Espèces Réactives de l'Oxygène) et une activation des caspases aboutissant à l'apoptose cellulaire (Bauer et al., 1999). Néanmoins, sa sur-expression dans des C2C12 ne provoque aucune toxicité (Bosnakovski et al., 2008a).

Une étude a mis en évidence que *ANT1* était sur-exprimé chez les patients FSHD (Gabellini et al., 2002) mais à son tour ces résultats furent infirmés par d'autres équipes (Jiang et al., 2003; Osborne et al., 2007). Tout comme *FRG2*, le modèle de souris sur-exprimant *ANT1* n'a montré aucun phénotype musculaire.

5.4 FAT1

L'organisation du muscle squelettique est une composante essentielle à sa fonction. Les myoblastes s'alignent progressivement afin d'initier la formation des myotubes puis des myofilaments. L'interaction cellule-cellule est donc fondamentale pour l'architecture multicellulaire. Un grand nombre de molécules participe à cette structure mais la famille des

cadhérines représente la principale. Parmi elle, le gène *FAT1* joue un rôle central dans la polarité planaire des myoblastes au cours du développement. Localisé en 4q35.2 (Dunne et al., 1995; Katoh and Katoh, 2006) ce gène de 131 kb est situé à 3,6 Mb des répétitions *D4Z4* et code pour protéine de 500 kDa localisée au niveau des leading edge de la cellule (Tanoue, 2005; Braun et al., 2007) mais aussi au niveau du DHPR dans le muscle. L'invalidation de ce gène entraîne une mortalité néonatale chez la souris (Ciani et al., 2003) tandis que la suppression des répétitions cadherines extracellulaires provoque une létalité à 50% (Caruso et al., 2013). Une souris sous-exprimant le gène *FAT1* présente des anomalies développementales et un phénotype assimilable à la FSHD (Caruso et al., 2013). La plupart des modèles transgéniques décrits précédemment ne développent que peu ou pas de phénotypes caractéristiques de la FSHD. C'est pourquoi ce modèle sous-exprimant le gène *FAT1* (gène impliqué dans la polarité planaire) est certainement de loin le plus intéressant d'un point de vue phénotypique. En effet ces souris dans lesquelles l'expression de *FAT1* a été réduite reproduisent à la fois les aspects musculaires et non-musculaires spécifiques du tableau clinique de la FSHD (Caruso et al., 2013). Au cours du développement, une carence en *FAT1* perturbe la migration des myoblastes tandis qu'aux stades postnataux, les souris mutantes pour *FAT1* présentent une réduction sélective de la masse musculaire, au niveau des épaules, suivies par une dystrophie musculaire plus étendue à des stades plus avancés. De manière intéressante, la sévérité des atteintes musculaires est directement reliée au niveau d'expression de *FAT1*. En outre ces souris présentent également des défauts non-musculaires caractéristiques de la FSHD tels qu'une vasculopathie rétinienne ou encore un défaut structurel au niveau de l'oreille caractéristique d'une surdité précoce. Ainsi ce modèle est le premier à récapituler les principaux phénotypes caractéristiques des patients FSHD.

Au regard de ces résultats chez la souris, la quantification de *FAT1* dans des biopsies fœtales FSHD 1 ou Contrôles a montré que celui-ci était spécifiquement sous-exprimé dans le muscle squelettique des fœtus FSHD (aucune différence dans le cerveau) ? .

6. Le modèle candidat : DUX4

La plupart des gènes étudiés ne font pas consensus quant à leurs implications dans la physiopathologie de la FSHD car l'existence de patients présentant de larges délétions du locus 4q35 a été décrite. Puisque l'absence de contractions n'est pas pathogénique, l'analyse des sous unités *D4Z4* a été approfondie.

6.1 Le motif D4Z4

Le séquençage complet du motif *D4Z4* a permis de mettre en évidence la présence de différents domaines au sein de chaque sous unité *D4Z4* (**Figure 20**). Ainsi, les répétitions *D4Z4* contiennent une répétition *LSau*, qui est un motif couramment observé dans des régions d'hétérochromatine et connu comme ayant un lien avec des phénomènes « d'effets de positions » (modification locale de la structure chromosomique pouvant moduler les niveaux d'expression de certains loci à proximité), ainsi qu'un motif *hhspm3* (« *Hypermethylated in Sperm* », duplicon riche en GC). La présence de ces motifs au sein de chaque répétition argumente l'hypothèse que les délétions *D4Z4* pourraient avoir un effet de positions et être impliquées dans le développement de la FSHD.

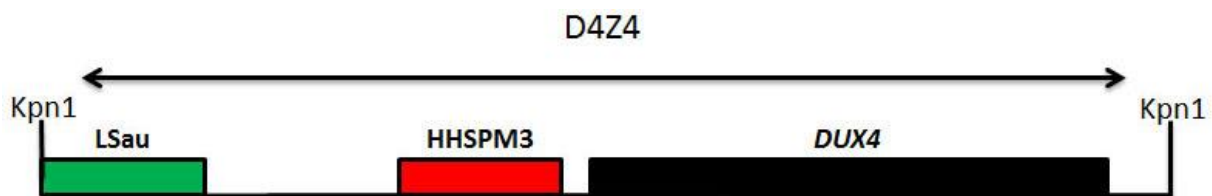


Figure 20 : Structure d'une répétition *D4Z4* : Le motif *D4Z4* de 3,3Kb est borné par les enzymes de restrictions Kpn1. Il contient un motif *LSau* (ntd 1-340), un motif *hhspm3* (ntd 1313-1780) et le gène *DUX4* (ntd1792-3063).

Le motif le plus intéressant découvert est un domaine à double homéoboîte contenue dans une phase ouverte de lecture (ORF) de 1245 pb appelé *DUX4* (van Geel et al., 1999). Les homéoboîtes sont des motifs de 180pb codant pour des domaines de 60 amino acides capables d'interagir avec l'ADN. En effet, ils codent pour trois hélices de type alpha qui interagissent via une séquence spécifique avec le grand sillon de l'ADN (Scott, 1992). Cette caractéristique est commune à toutes les protéines appartenant à la famille des facteurs de

transcriptions. Chaque sous unité *D4Z4* possède l'ORF *DUX4* mais il semblerait que seule la répétition *D4Z4* la plus proche des télomères soit capable de produire un ARNm *DUX4* stable (Lemmers et al., 2010). En effet, dans un contexte de chromosome permissif 4qA, le site de polyadénylation, absent du chromosome 4qB, permet de stabiliser l'ARNm. En cas d'absence de celui-ci (haplotype 4qB) l'ARNm est dégradé.

Puisque les chromosomes 4 et 10 présentent 98% d'homologie pour les répétitions *D4Z4*, les ORF de *DUX4* sont présentes sur les 2 chromosomes. De ce fait, la production d'un ARNm *DUX4* peut avoir lieu à partir des 2 chromosomes. Néanmoins, bien que le chromosome 10 possède également l'haplotype 4qA permissif, un SNP (Single Nucleotide Polymorphism) sur le signal de poly(A) a été décrit, transformant le ATTAAA canonique en ATCAAA ou ATTTAA, non reconnu comme signal de polyadénylation (Lemmers et al., 2010) (**Figure 21**). Ainsi le transcrit *DUX4* du chromosome 10 ne bénéficie pas de l'effet stabilisant d'un signal de polyadénylation pouvant expliquer que les contractions *D4Z4* issues de ce chromosome ne soient pas pathogéniques.

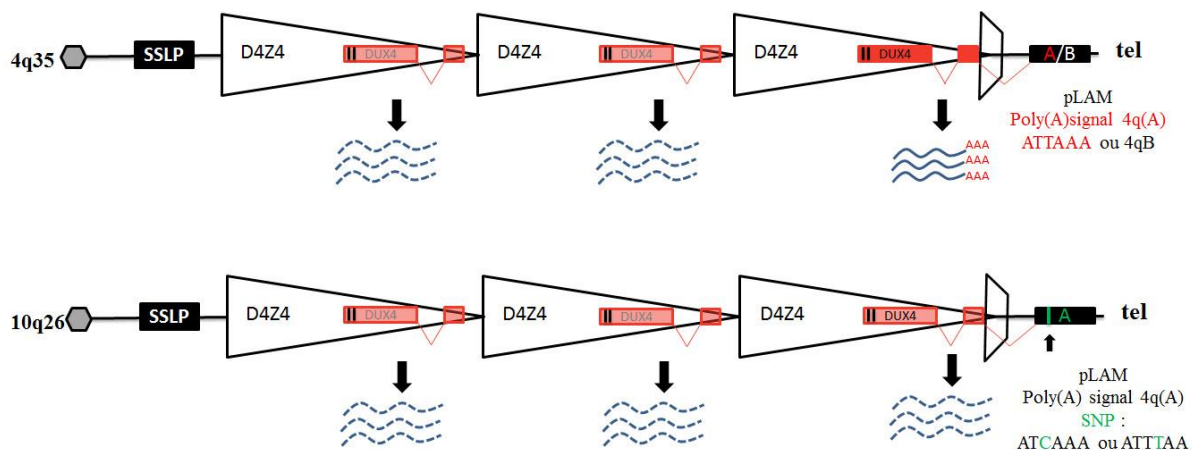


Figure 21 : Représentation de l'ORF *DUX4* présent dans les répétitions *D4Z4* du chromosome 4q et 10q (adapté de Lemmers et al., 2010) : Les chromosomes 4 et 10 présentent 98% d'homologie pour les répétitions *D4Z4*, les ORF de *DUX4* sont présentes sur les deux chromosomes. Seule la dernière répétition *D4Z4* serait capable de produire un ARNm *DUX4* stable, dans un contexte 4qA, de par sa proximité avec le signal de poly(A). Un SNP (en vert) sur le signal de poly(A) du chromosome 10 le rend non fonctionnel empêchant ainsi la production d'un ARNm *DUX4* stable expliquant pourquoi les contractions sur le chromosome 10 ne sont pas pathogéniques.

6.2 DUX4 : un facteur de transcription

6.2.1 Les différents transcrits identifiés

Différents variants de l'ARNm *DUX4* ont été mis en évidence (Dixit et al., 2007; Snider et al., 2009). En effet, un épissage alternatif est retrouvé dans la région 3'UTR (Untranslated Region) localisé dans pLAM et retrouvé uniquement en aval de la dernière répétition *D4Z4* (**Figure 22**). Ainsi, 2 variants longs, appelés *DUX4-FL* (full length), ont été caractérisés (Snider et al., 2010). Ils ne se différencient que par l'épissage alternatif du premier intron (136pb) tandis que le second est constitutivement épissé (365 pb). Malgré cet épissage les 2 isoformes donnent naissance à la même protéine car les régions 3'UTR ne font pas partie de la phase codante. Ces deux isoformes ont été détectées simultanément dans des cellules FSHD (Ferreboeuf et al., 2013) mais nous ne sommes pas en mesure de définir si elles coexistent au sein de la même cellule. Le rôle de cet épissage reste inconnu mais il pourrait jouer sur la stabilité du transcrit.

D'autres variants longs ont également été identifiés et sont retrouvés dans la lignée germinale (testicules). En effet, l'exon 3 (contenant le signal de polyadénylation) est épissé et rejoint les exons 4 à 7 contenant un second signal de polyadénylation localisé dans l'exon 7 (Snider et al., 2010) et cela indépendamment d'un génotype FSHD. Dans les testicules, les transcrits produits sont issus à la fois du chromosome 4 et du chromosome 10 en égale quantité (Snider et al., 2010).

Bien que les deux signaux de poly(A) (exon 3 et 7) soient utilisables, il semblerait que le signal de l'exon 3 soit préférentiellement utilisé (Snider et al., 2010). De plus, l'haplotype 4qB ne contenant pas cette région ne possède pas ces signaux alternatifs confirmant bien son statut de non permissif.

Une troisième isoforme courte, appelée *DUX4-S* (Small), a également été identifiée (Snider et al., 2010). Dans ce variant, un site cryptique d'épissage localisé dans l'ORF de *DUX4* est utilisé. Le site accepteur d'épissage est celui de l'intron 1, ce qui conduirait à une protéine tronquée conservant son domaine de fixation à l'ADN (double homéobox) mais dépourvue de la région C terminale portant le domaine actif.

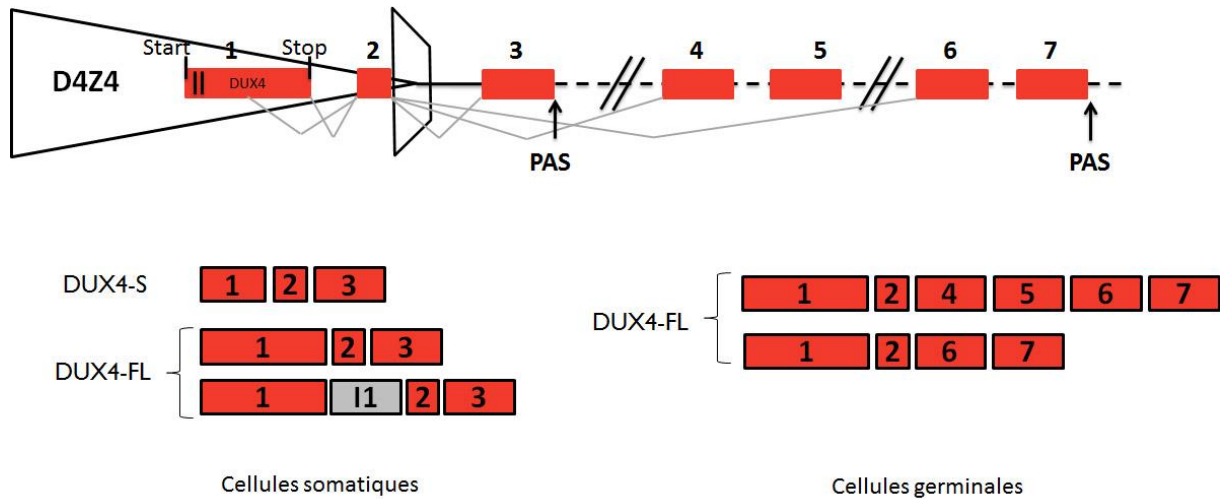


Figure 22 : Schéma récapitulant les isoformes de *DUX4* produites selon les cellules (adapté de Snider et al., 2010) : La forme *DUX4-S* est produite chez les individus sains par utilisation d'un site cryptique d'épissage situé dans l'ORF tronquant alors la protéine de son domaine transactivateur. Dans la région 3'UTR, le premier intron est épissé de manière alternative générant deux isoformes *DUX4-FL* dans les cellules somatiques, tandis que le second intron est épissé de manière constitutive. Dans les cellules germinales, l'exon 3 (contenant le signal de poly(A)) est épissé pour générer des isoformes contenant des exons localisés plus en aval et utiliser un signal de poly(A) alternatif localisé dans l'exon 7.

6.2.2 Expression chez les patients FSHD

L'identification de 3 isoformes de *DUX4*, dont une non fonctionnelle, liée à l'utilisation d'un site cryptique d'épissage localisé dans l'ORF, a ouvert une nouvelle voie d'étude. La présence de ces différents transcrits a été recherchée parmi un panel de cellules et biopsies musculaires provenant d'individus adultes Contrôles (allèle permissif) ou FSHD1-2. Dans une première étude, les formes *DUX4-FL* ont été détectées exclusivement dans les échantillons FSHD (dans 50% des biopsies, 100% dans les cellules). L'isoforme *DUX4-S* a quant à elle été détectée dans des biopsies (50%) et cellules contrôles (100%) mais également dans les biopsies et cellules FSHD mais à une moindre fréquence (40 et 30%) (Snider et al., 2010). De manière intéressante, *DUX4-FL* est retrouvé exprimé de manière constitutive dans d'autres organes, notamment dans la lignée germinale (testicules). Une seconde étude portant sur un plus grand nombre d'échantillons contrôles et FSHD a confirmé la majorité de ces résultats (Jones et al., 2012). En effet, *DUX4-FL* est retrouvée dans 78% des biopsies et

cellules FSHD mais de manière intéressante, l'isoforme longue est aussi détectée dans 12% des échantillons contrôles possédant un allèle permissif. Ainsi, bien que *DUX4-FL* soit exprimée de manière consistante et quasi systématique chez les individus FSHD (Snider et al., 2010; Tsumagari et al., 2011; Jones et al., 2012), elle n'est pas exclusive aux patients (**Figure 23**).

D'après Snider et al., 2010	<i>DUX4-FL</i>		<i>DUX4-S</i>	
	FSHD	Contrôles	FSHD	Contrôles
Cellules	100%	nd	30%	100%
Biopsies	50%	nd	40%	50%
D'après Jones et al., 2013	<i>DUX4-FL</i>		<i>DUX4-S</i>	
	FSHD	Contrôles	FSHD	Contrôles
Cellules	92%	44%	-	-
Biopsies	78%	12%	-	-

Figure 23 : Récapitulatif des fréquences de détection des isoformes de *DUX4* publié :

nd : non détecté

L'étude de l'expression de *DUX4* a été également réalisée sur des fibroblastes FSHD et Contrôles reprogrammés en cellules iPS afin de déterminer si *DUX4* peut être exprimé naturellement dans des états très précoces de la cellule (Snider et al., 2010). Comme attendu, les fibroblastes FSHD seuls, reprogrammés en iPS ou différenciés en corps embryonnaire, montrent une expression de *DUX4-FL*. Les fibroblastes Contrôles expriment *DUX4-S*, or une fois reprogrammés, ces cellules expriment *DUX4-FL* qui disparaît lorsqu'elles se différencient en corps embryonnaire. Ce changement d'isoformes peut s'expliquer par une forte diminution des marques d'hétérochromatine (triméthylation des H3K9) entre les iPS et les fibroblastes contrôles dont ils sont issus. Il apparaît donc que l'expression des différentes isoformes est soumise à une empreinte épigénétique.

6.2.3 Un modèle d'expression pour DUX4

Des mesures de transcrits et des immunofluorescences ont mis en évidence la très faible expression de DUX4 chez les individus FSHD. Ainsi, la fréquence d'apparition de DUX4 est d'environ 1 noyau sur 1000. Malgré cela, son effet sur un grand nombre de gènes cibles est lui très fortement prononcé et quantifiable (Dixit et al., 2007; Geng et al., 2011, 2012; Snider et al., 2010). Dans ces conditions, comment une protéine si faiblement exprimée peut-elle avoir un effet global ? Deux modèles d'expression pour DUX4 ont alors été émis (Ehrlich and Lacey, 2012) (**Figure 24**). Le premier, appelé « Majority Rules », postule sur une très faible expression de DUX4 dans un grand nombre de noyaux mais selon une cinétique de temps non synchronisée. Ainsi, chaque noyau exprimant ou ayant exprimé DUX4 montre des dérégulations géniques majeures persistantes alors même que DUX4 n'est plus détecté. Le second modèle quant à lui, appelé « Minority Rules », table sur une expression dans un pool très réduit de noyaux suivis d'une diffusion de la protéine et donc du phénomène pathologique.

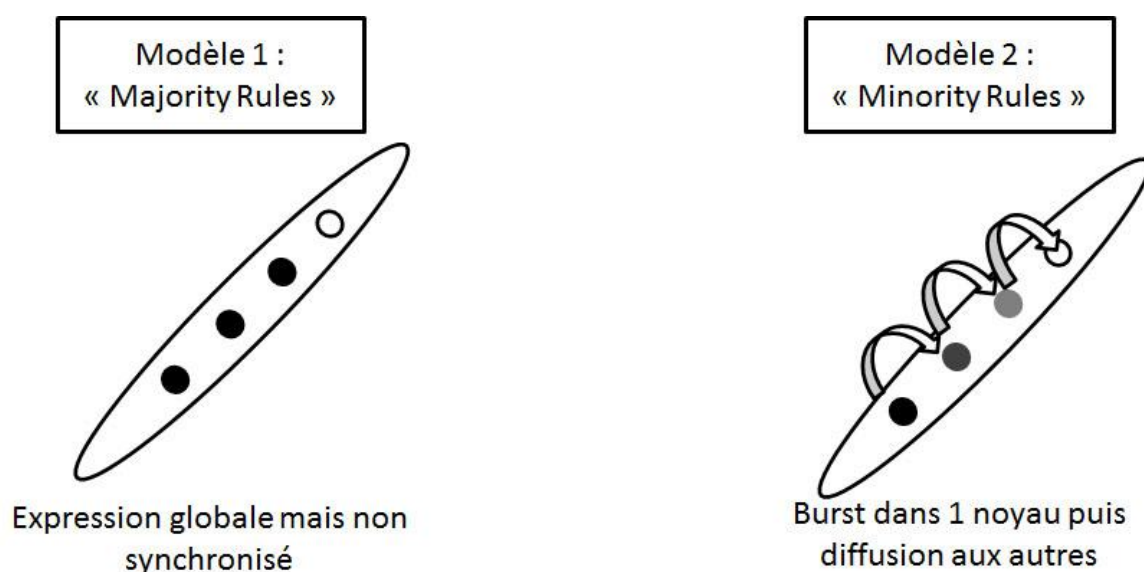


Figure 24 : Modèle d'expression pour DUX4 (adapté de Ehrlich et al., 2012) : Le premier modèle suppose que *DUX4* est exprimé dans un grand nombre de noyaux mais que ceux-ci n'étant pas synchronisé, sa détection par immunofluorescence ne permet d'observer qu'un nombre limité de noyaux. Dans le second modèle, *DUX4* serait transcrit dans un nombre limité de noyaux mais une fois traduite, la protéine serait à même de diffuser d'un noyau à l'autre.

6.3 Conséquences liées à l'expression de DUX4

6.3.1 Toxicité et dérégulations *in vitro*

Puisque DUX4 peut jouer le rôle de facteur de transcription (Dixit et al., 2007; Geng et al., 2012) sa dérégulation chez les patients FSHD pourrait induire des défauts phénotypiques majeurs. Sa forte toxicité a d'ailleurs été démontrée lors d'essais de sur-expression *in-vitro* (Kowaljow et al., 2007; Bosnakovski et al., 2008a; Wuebbles et al., 2010; Wallace et al., 2011; Tassin et al., 2013). La sur-expression de DUX4 provoque une activation des caspases 3, 6, 8 et 9, (marqueurs de l'apoptose) ainsi que des changements phénotypiques liés à l'induction de la voie apoptotique (ralentissement de la prolifération, arrondissement des cellules puis mort). De manière intéressante, l'effet toxique de DUX4 est amoindri lorsque son expression est induite au stade myotube suggérant un effet toxique lié à une possible interférence avec des marqueurs myogéniques. En effet, l'expression de DUX4 induit une diminution d'un certain nombre de marqueurs myogéniques comme la myogénine, la desmine, MYOD provoquant une altération de l'efficacité de fusion (Bosnakovski et al., 2008a). De plus, l'utilisation d'antioxydants associés à la sur-expression de DUX4 assure une diminution significative de la mortalité confirmant que DUX4 induit une plus forte sensibilité des cellules au stress oxydant. Ces résultats sont ainsi en accord avec une étude protéomique sur des biopsies FSHD montrant que les protéines de réponse au stress oxydatif, en partie régulé par MYOD, et MYOD lui-même, sont présentes en moins grandes quantités comparées à des biopsies contrôles (Celegato et al., 2006).

L'homéodomaine de DUX4 est très fortement similaire à celui retrouvé sur les facteurs de transcription PAX3 et PAX7 (exprimé dans les cellules satellites et au stade myoblaste). Il a ainsi été démontré que DUX4 entre en compétition avec PAX3/PAX7 et qu'une surexpression des PAX sauve la cellule de l'effet toxique de DUX4 (Bosnakovski et al., 2008a) (**Figure 25**).

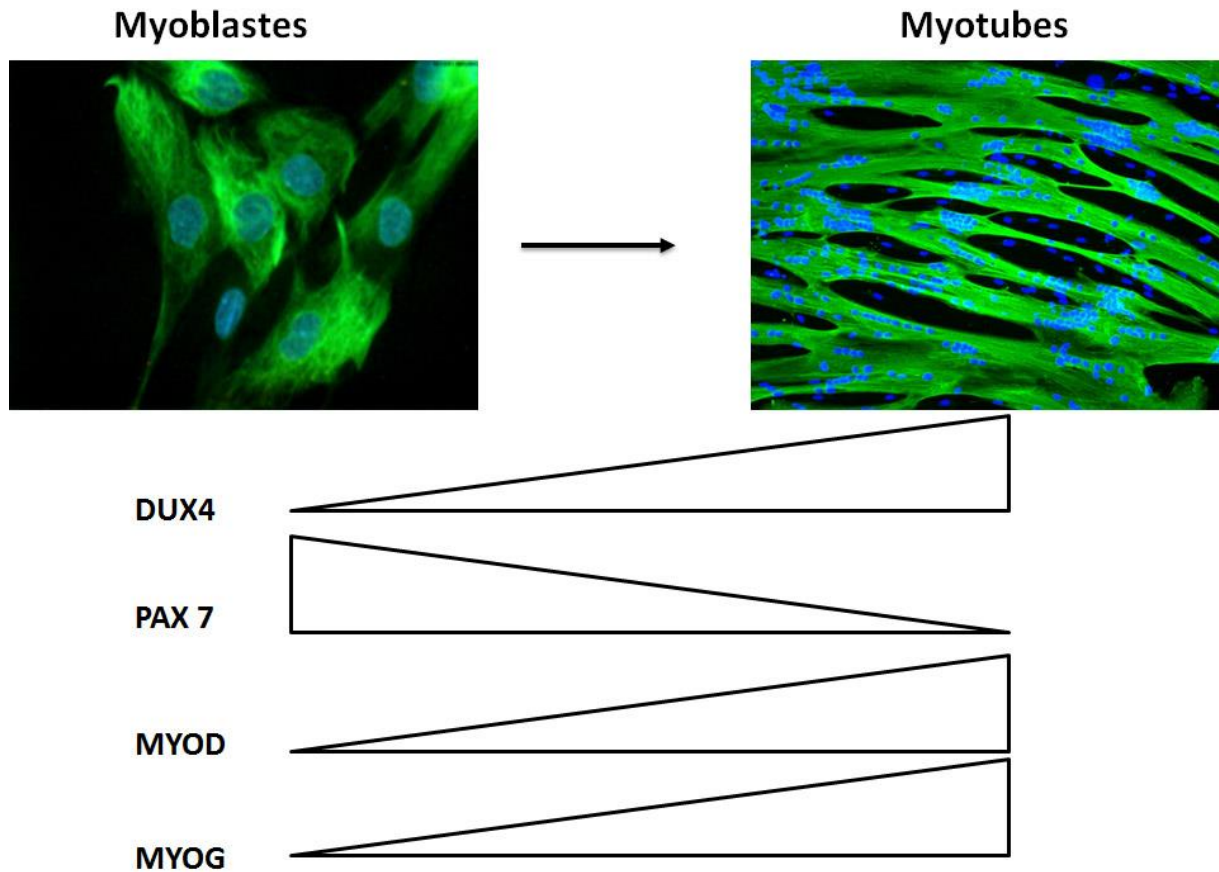


Figure 25 : Expression de DUX4 et des marqueurs myogéniques au cours de la différenciation : *DUX4* est progressivement exprimé au cours de la différenciation tout comme les marqueurs myogéniques MyoD et MyoG tandis que les marqueurs de cellules satellites PAX7 décroissent rapidement après l'activation des myoblastes. DUX4 serait capable d'interférer directement ou indirectement avec ces différents marqueurs pouvant alors altérer la mise en place du programme myogénique et la différenciation cellulaire.

Puisque DUX4 appartient à la famille des doubles homéoboxes, il peut jouer le rôle de facteur de transcription. La comparaison du profil d'expression sur le génome entier, entre des biopsies FSHD et des biopsies provenant d'autres pathologies neuromusculaires, a mis en évidence une sur-expression de 5 gènes, notamment d'un facteur à double homéobox appelé PITX1. La fusion du promoteur de *PITX1* à un gène rapporteur a mis en évidence que DUX4 est capable de se fixer à une séquence spécifique de 30 pb sur le promoteur et d'activer ainsi l'expression du rapporteur (Dixit et al., 2007). La mutation du motif TAAT abolit la transactivation.

Afin d'identifier des cibles directement régulées par DUX4, une construction permettant la sur-expression des isoformes de DUX4 (DUX4-FL et DUX4-S) a été introduite

par lentivecteur dans des cellules musculaires primaires contrôles (Geng et al., 2012). Puisque DUX4 est toxique pour les cellules, l'analyse par microarray a été réalisée 24H après la transduction. DUX4-FL induit la sur-expression de 466 gènes et la sous-expression de 244 gènes tandis que DUX4-S ne provoque qu'une faible modulation (37 gènes sur-exprimé et 1 seul sous-exprimé). Le panel des modulations induites par DUX4-FL est très large, il regroupe notamment des gènes retrouvés dans les cellules souches et la lignée germinale, le processing des ARNs, la voie ubiquitine ou encore la défense immunitaire innée. De plus parmi les gènes induits par DUX4-FL, un grand nombre ne sont normalement pas exprimés dans les myoblastes. La comparaison de cellules et biopsies FSHD avec des Contrôles a confirmé une up-régulation d'une grande partie des gènes dérégulés lors d'une sur-expression de DUX4-FL. Des essais de Si-RNA sur DUX4-FL ont confirmé la spécificité de transactivation médiée par DUX4-FL et pas par DUX4-S (dominant négatif).

DUX4 est retrouvé exprimé dans la lignée germinale (testicules) dans laquelle aucune toxicité n'est observée (Snider et al., 2010). Puisque DUX4 induit l'expression de gènes liés à la lignée germinale, il est donc possible que son expression anormale induise une toxicité liée à une activation d'un programme gamétogénique, incompatible avec les cellules du muscle squelettique.

6.3.2 Toxicité et dérégulations *In vivo*

D'un point de vue développemental, l'expression de DUX4-FL dans le zebrafish a confirmé son impact sur l'expression de la myogénine ou de MyoD (Mitsubishi et al., 2012), provoque une forte mortalité des embryons et induit des défauts phénotypiques majeurs comme la forme du corps (Wallace et al., 2011; Mitsubishi et al., 2012). En revanche, l'expression de DUX4-S ou de DUX4 muté dans son domaine de fixation à l'ADN n'induit pas ces défauts. Enfin, la co-injection de DUX4-FL et de DUX4-S montre une compétition entre les deux isoformes. Ainsi, DUX4-S se comporte comme un dominant négatif de DUX4-FL, il bloque les sites de fixation pour DUX4-FL sur l'ADN, et interfère donc avec l'effet toxique de DUX4-FL médié par la transactivation de ses gènes cibles (Geng et al., 2012; Mitsubishi et al., 2012) (**Figure 26**).

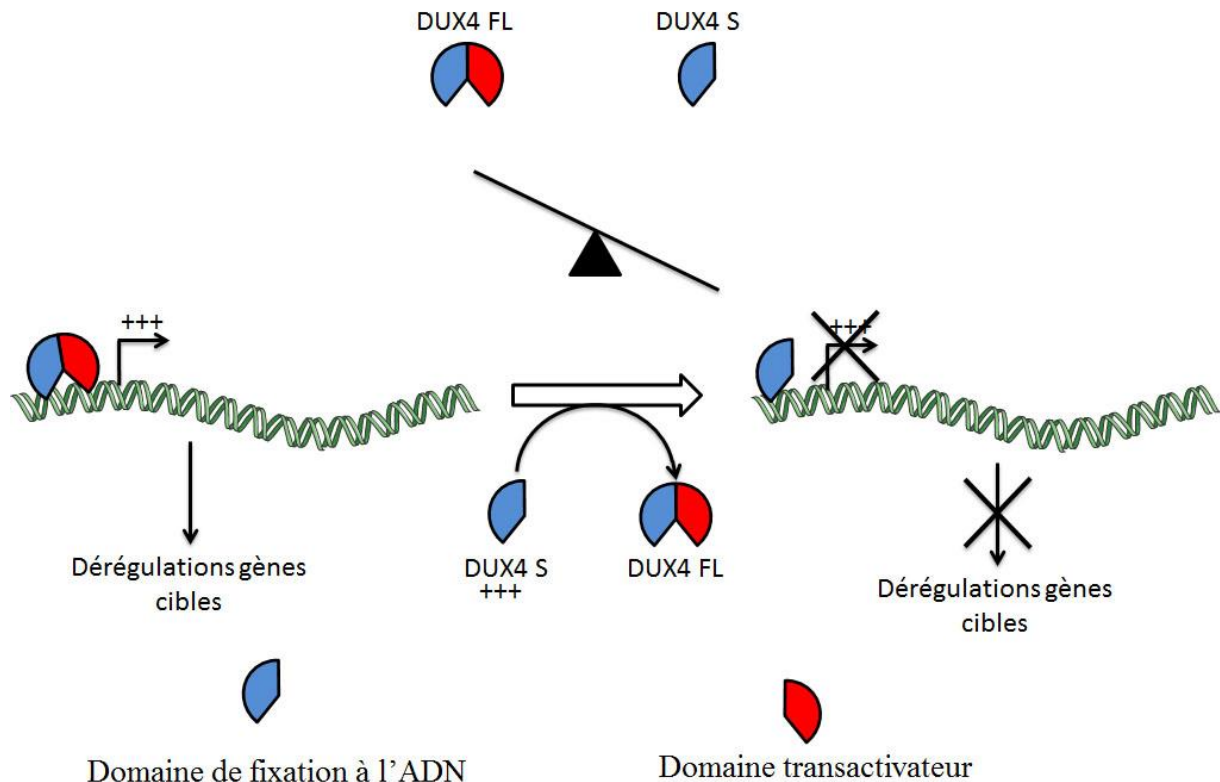


Figure 26 : Modélisation de l'effet de compétition entre DUX4-FL et DUX4-S pour leurs cibles: DUX4-FL possède deux domaines de fixation à l'ADN (domaine HOX) ainsi qu'un domaine transactivateur permettant à la protéine de jouer sa fonction de facteur de transcription. DUX4-S ne possède que la région N-terminale lui permettant uniquement de se fixer à l'ADN sans pouvoir jouer de rôle transactivateur. Ainsi, DUX4-S peut entrer en compétition pour les sites de fixations à l'ADN avec DUX4-FL et jouer un rôle de dominant négatif.

In Vivo, l'injection d'un AAV6-DUX4 dans le tibialis antérieur d'une souris sauvage provoque de graves lésions phénotypiques (dégénérescence) médiées par la voie p53 et l'augmentation des caspases (voie apoptotique). Ces lésions ne sont pas présentes lorsque le domaine de fixation à l'ADN est muté sur DUX4 (Wallace et al., 2011) confirmant ainsi *in vivo* que la surexpression de DUX4 est toxique pour le tissu musculaire.

Un modèle de souris portant les répétitions *D4Z4* a été généré afin d'étudier *In-vivo* le mécanisme physiopathologique lié à DUX4. Une première lignée de souris contient un fragment EcoR1 de 2,5 répétitions *D4Z4* provenant d'un allèle FSHD 1 permissif (4qA161 et PAS) et la seconde un fragment EcoR1 de 12,5 répétitions provenant lui aussi d'un allèle permissif (Krom et al., 2013). *DUX4* est ainsi plus fortement exprimé dans la souris 2,5 *D4Z4*

et diminue progressivement au cours du développement mais reste supérieur à la souris 12,5. Il est retrouvé exprimé dans une majorité de tissus squelettiques (cœur, diaphragme, quadriceps etc...) mais également dans une variété d'autres tissus non musculaires (yeux, testicules, cerveau, foie etc...). Enfin, comme décrit, le niveau de *DUX4* augmente au cours de la différenciation musculaire (myoblastes versus myotubes).

D'un point de vue épigénétique, la souris 2,5 montre un niveau de méthylation plus faible que la souris 12,5 en accord avec l'impact des contractions *D4Z4* sur la méthylation de la région 4q chez les patients FSHD 1. De plus, un certain nombre de gènes cibles de *DUX4* ont été identifiés comme dérégulés chez la souris 2,5 confirmant la fonctionnalité de *DUX4* dans ce modèle. Néanmoins, d'un point de vue phénotypique, la souris 2,5 n'a développé aucune altération musculaire caractéristique. Environ 54% des souris 2,5 développent une anomalie de développement des yeux d'origine inconnue.

Bien que ce modèle récapitule plusieurs aspects de la FSHD comme les contractions, l'hypométhylation et l'expression de *DUX4*, l'absence de phénotypes musculaires tend à prouver que les contractions, aboutissant à l'expression de *DUX4*, ne sont pas capables à elles seules de conduire au développement de la FSHD.

Enfin, une xénogreffe de biopsies musculaires provenant de familles non atteintes ou de patients FSHD, a été réalisée chez une souris immunodéficiente afin de suivre In-vivo le comportement de ces muscles. Le tibialis antérieur de la souris est remplacé par la biopsie humaine et ligaturé aux tendons proximaux et distaux. Spontanément, la biopsie se vascularise, se régénère et exprime les marqueurs membranaires humains comme la spectrine ou la Lamin A/C. Des mesures de contractibilité (mesure du couplage excitation-contraction lié à une impulsion électrique) et de forces sont en adéquation avec un muscle fonctionnel. La présence de *DUX4-FL* est confirmée dans les xénogreffes de muscles FSHD et les dérégulations des gènes cibles sont concordantes avec ce qui est observé entre les biopsies FSHD et Contrôles d'origines. Néanmoins, aucun phénotype dystrophique n'est observé dans le muscle greffé (provenant pourtant de quadriceps FSHD moyennement atteint) suggérant que l'environnement musculaire joue un rôle dans la pathologie. Il est également important de noter que les souris utilisées pour la xénogreffe sont immunodéficientes il est donc possible d'envisager un rôle du système immunitaire dans la physiopathologie de la FSHD.

Malgré ces preuves de principe sur la toxicité de *DUX4*, aucune différence phénotypiques majeures (apoptose, défauts de différenciation) n'ont été clairement identifiées lors de comparaisons de cultures cellulaires d'individus FSHD et Contrôles hormis une

réponse accrue au stress oxydant (Barro et al., 2010; Turki et al., 2012). De plus, DUX4 ne soit pas ou peu exprimé au stade myoblastes et n'interfère donc pas (ou très faiblement) avec certains marqueurs myogéniques (PAX3/7).

Ainsi, il est probable que la très faible expression de DUX4 *in vivo* ne soit suffisante pour induire la toxicité observée par sur-expression.

Au vu des études réalisées, la FSHD apparaît comme une pathologie où les aspects épigénétiques sont prédominants. Il apparaît également que la présence du motif *D4Z4* est déterminante dans le développement de la pathologie. Les contractions *D4Z4* seraient ainsi à l'origine de la modification du profil épigénétique du chromosome 4q ainsi que de phénomènes de localisation chromosomique pouvant induire des dérégulations géniques en Cis et en Trans. L'existence des patients FSHD 2, où les contractions *D4Z4* ne sont pas présentes mais où une forte hypométhylation de la région est retrouvée, accredit la thèse épigénétique. Puisque l'absence de contractions *D4Z4* n'est pas pathogénique, il apparaît que la production du facteur de transcription DUX4 à partir des répétitions est indispensable à l'apparition de la pathologie. L'identification d'un grand nombre de gènes cibles dérégulés directement par DUX4 confirme que son expression anormale chez les patients peut entraîner des cascades de dérégulations, néanmoins, aucun lien n'a encore été établi entre les dérégulations de ces gènes cibles et la physiopathologie de la FSHD. Ainsi, ces gènes cibles peuvent être utilisés comme des biomarqueurs de l'expression de DUX4. Puisque l'origine développementale de la FSHD a été suggérée par certaines études (Richards et al., 2012; Snider et al., 2010) et qu'il apparaît que le gène *FAT1*, dont la sous expression chez la souris induit les phénotypes musculaires et non musculaires caractéristiques de la FSHD, est spécifiquement sous-exprimé dans les muscles de fœtus FSHD, mon travail a consisté à caractériser l'expression de DUX4 au stade fœtal ainsi qu'à identifier si les marqueurs moléculaires de la FSHD sont retrouvés à ce stade de développement.

RESULTATS

OBJECTIFS DE LA THESE

La dystrophie Facio-Scapulo-Humérale est une pathologie du muscle squelettique touchant de manière progressive et asymétrique les muscles du visage, des épaules et des bras. Bien qu'un grand nombre d'études ait été réalisé afin d'identifier les causes moléculaires de la pathologie, aucun gène candidat ne fait consensus hormis l'expression anormale d'une protéine appelée DUX4. Sa fonction de facteur de transcription le place comme étant un acteur clé du développement de la pathologie. De plus, sa toxicité dans les cellules mononuclées a été clairement démontrée, néanmoins, son très faible niveau d'expression in-vivo ainsi que l'absence de signes évidents de surmortalité in-vitro remettent en cause l'aspect toxique (du moins dans des conditions d'expression normale).

Bien que le défaut génétique soit présent de manière constitutive chez la majorité des patients FSHD, les premiers signes cliniques n'apparaissent que durant la seconde décade. Cependant, un petit nombre de patients, portant les contractions les plus grandes, développent une forme infantile très grave.

La FSHD étant décrite comme une maladie du jeune adulte, la majorité des études ont été réalisées à ce stade. Or la recherche de gènes candidats, notamment musculaire, est délicate à ce stade car le muscle squelettique est déjà endommagé (changement de la taille des fibres, infiltration adipocytaire, tissus fibrotique et inflammatoire), il est ainsi compliqué de déterminer les facteurs primaires ayant provoqué la cascade de dérégulation.

C'est pourquoi durant mes trois années de thèse, je me suis intéressé à caractériser les voies moléculaires dérégulées chez des fœtus porteurs ou non du génotype FSHD. Le rôle de DUX4 dans la physiopathologie de la FSHD n'étant à ce jour pas clairement établi, j'ai dans un premier temps mis en évidence l'expression de DUX4-FL dans des fœtus FSHD. Dès lors, mes travaux ont consisté à déterminer si les cascades de dérégulations moléculaires observées au stade adulte étaient présentes également à un stade très précoce en l'absence de tout dommage musculaire.

PUBLICATION N°1: DUX4 and DUX4 downstream target genes are expressed in fetal FSHD muscles

Ferreboeuf M, Mariot V, Bessi res B, Vasiljevic A, Atti -Bitach T, Collardeau S, Morere J, Roche S, Magdinier F, Robin-Ducellier J, Rameau P, Whalen S, Desnuelle C, Sacconi S, Mouly V, Butler-Browne G, Dumonceaux J. Hum Mol Genet. 2013

R sum 

La FSHD est une maladie progressive qui se caract rise par une diff rence d'atteinte entre les muscles. Ainsi, le trap ze et le rhombo ide font partie des muscles atteints les plus pr cocement contrairement au quadriceps ou au sous-scapulaire. DUX4  tant l'un des g nes candidats le plus pl biscit , nous avons regard  si son expression  tait d tect e dans des cellules f tales d riv es d'un quadriceps et d'un trap ze FSHD (m me f tus de 16A). A ce stade de d veloppement, les muscles ne sont pas consid r s comme atteints et ne pr sentent donc aucune marque dystrophique.

Nos travaux ont permis de mettre en  vidence pour la premi re fois que DUX4 est bien exprim  dans des cellules f tales FSHD. De plus son niveau d'expression est plus fort dans le quadriceps compar  au trap ze provenant du m me f tus. N anmoins, la quantification de diff rents marqueurs myog niques comme la Myog nine et l'inclusion de l'exon 11 de BIN1 montrent une diff rence dans l'initiation du programme myog nique entre les 2 muscles. Nous avons  galement pu constater que les g nes cibles de DUX4 sont aussi d r gul s dans les cellules f tales avec une cin tique d'expression superposant celle de DUX4. L'analyse sur des biopsies f tales FSHD et Contr les montre une tr s grande variabilit  inter-individus, ainsi, seulement 2 g nes valid s en culture cellulaire sont retrouv s sp cifiquement surexprim s faisant d'eux de bons marqueurs directement reli s   l'expression de DUX4. La comparaison du niveau d'expression de ces 2 g nes entre des biopsies adultes et f tales FSHD est similaire sugg rant que le niveau d'expression de DUX4 n'est pas diff rent chez l'adulte.

Pris ensemble, ces r sultats sugg rent que, malgr  l'apparition tardive des sympt mes, les premiers d fauts mol culaires sont pr sents d s le stade f tal.

DUX4 and DUX4 downstream target genes are expressed in fetal FSHD muscles

Maxime Ferreboeuf^{1,†}, Virginie Mariot^{1,†}, Bettina Bessi eres^{2,3}, Alexandre Vasiljevic⁴, Tania Atti e-Bitach^{2,3}, Sophie Collardeau⁴, Julia Morere⁵, St ephane Roche⁵, Fr ed erique Magdinier⁵, J er ome Robin-Ducellier⁶, Philippe Rameau⁷, Sandra Whalen⁸, Claude Desnuelle^{9,10}, Sabrina Sacconi^{9,10}, Vincent Mouly¹, Gillian Butler-Browne¹ and Julie Dumonceaux^{1,*}

¹INSERM U974, UMR 7215 CNRS, Institut de Myologie, UM 76 Universit  Pierre et Marie Curie, Paris 75013, France
²D epartement de G n tique, INSERM U781 et Fondation IMAGINE, Assistance Publique-H pitaux de Paris, H pital Necker-Enfants Malades, Paris 75015, France ³Universit  Paris Descartes, Paris 75006, France ⁴Centre de pathologie Est, CHU-Lyon, 59 bd Pinel, Bron 69677, France ⁵UMR_S910, INSERM-Aix-Marseille Universit , Facult  de M decine de la Timone, 27, Bd Jean Moulin, Marseille cedex 05 13385, France ⁶Department of Cell Biology, UT Southwestern Medical Center, Dallas, TX, USA ⁷Institut Gustave Roussy, Plate-Forme Imagerie et Cytom trie, 114 rue Edouard Vaillant, Villejuif 94805, France ⁸D epartement de g n tique, Groupe hospitalier Piti -Salp tri re, Paris 75013, France ⁹Centre de R f rence des Maladies Neuromusculaires and ¹⁰Centre National de la Recherche Scientifique (CNRS) Unit  Mixte de Recherche (UMR) 7277, Nice University Hospital, Nice, France

Received June 25, 2013; Revised August 13, 2013; Accepted August 14, 2013

Facioscapulohumeral muscular dystrophy (FSHD) is one of the most prevalent adult muscular dystrophies. The common clinical signs usually appear during the second decade of life but when the first molecular dysregulations occur is still unknown. Our aim was to determine whether molecular dysregulations can be identified during FSHD fetal muscle development. We compared muscle biopsies derived from FSHD1 fetuses and the cells derived from some of these biopsies with biopsies and cells derived from control fetuses. We mainly focus on DUX4 isoform expression because the expression of DUX4 has been confirmed in both FSHD cells and biopsies by several laboratories. We measured DUX4 isoform expression by using qRT-PCR in fetal FSHD1 myotubes treated or not with an shRNA directed against DUX4 mRNA. We also analyzed DUX4 downstream target gene expression in myotubes and fetal or adult FSHD1 and control quadriceps biopsies. We show that both DUX4-FL isoforms are already expressed in FSHD1 myotubes. Interestingly, DUX4-FL expression level is much lower in trapezius than in quadriceps myotubes, which is confirmed by the level of expression of DUX4 downstream genes. We observed that TRIM43 and MBD3L2 are already overexpressed in FSHD1 fetal quadriceps biopsies, at similar levels to those observed in adult FSHD1 quadriceps biopsies. These results indicate that molecular markers of the disease are already expressed during fetal life, thus opening a new field of investigation for mechanisms leading to FSHD.

INTRODUCTION

Facioscapulohumeral muscular dystrophy (FSHD) is one of the most frequent muscular dystrophies with an incidence of 4/100 000 (www.orpha.net, June 2013). The pathology often

begins during late adolescence with an asymmetric atrophy of muscles located in the face, the shoulder and in the arms. A high inter-individual variability is observed with very variable phenotypes being reported for the same genotype (1,2). In general, the pathology progresses slowly. The genetic cause of

the disease was proposed 20 years ago (3): FSHD is classically associated with the contraction of a macrosatellite repeat array in the sub-telomeric region of chromosome 4q35 (for review see (4)). This region normally consists of 11–100 tandem repeats of a 3.3 kb D4Z4 unit, and deletions leaving only 1 to 10 D4Z4 repeats have been associated with FSHD1 patients. In FSHD2 patients, the contraction of the D4Z4 array is not observed although the clinical features are identical to those observed in FSHD1 patients (5).

D4Z4 repeats are not restricted to chromosome 4, but FSHD1 has only been associated with D4Z4 contractions occurring on chromosome 4. Moreover, a complex genomic context (for review see (6)) is required for FSHD onset. Two allelic variants of chromosome 4q (4qA and 4qB) exist in the region distal to D4Z4 and FSHD seems to be almost always associated with 4qA (7–11), which contains a polyadenylation signal for transcripts encoded from the D4Z4 unit (12). Each D4Z4 region contains one open reading frame encoding the double homobox gene DUX4 (13), but owing to the presence of the polyadenylation signal located immediately after the D4Z4 repeats, only the last D4Z4 unit is able to generate a stable mRNA (12,14). In most of the cases, full-length DUX4 is expressed in FSHD patients whereas it is not expressed in control individuals (15–17). The expression of DUX4, always associated with the 4qA variant, has introduced the notion of a permissive chromosome for FSHD (12). Combined with the fact that at least one D4Z4 unit is required to develop FSHD (18), it has been proposed that DUX4 may play a role in the development of the pathology, although its role is still unclear. The pre-messenger RNA can generate at least three different mRNAs varying in size (15). The shortest one (DUX4-S) can be found in both control and FSHD biopsies as well as in cell culture, whereas the largest one (DUX4-FL) is mainly found in FSHD individuals (16,17). The two DUX4-FL isoforms differ by the retention or not of the first intron located in the 3'UTR. Overexpression of DUX4-FL protein appears to be highly toxic (19–22), and its expression has been proposed to generate a disorganization of the nuclear envelope (23), atrophic myotubes (20), an activation of the PITX1 gene (14), but also of germline genes, retroelements and immune mediators (24), and may cause a p53-dependent myopathy in vivo (25). Interestingly, DUX4 expression has also been described in FSHD2 patients (16) and these patients also carry the 4qA haplotype (12).

In FSHD1, D4Z4 contraction has been associated with more relaxed chromatin structure leading to the expression of DUX4. FSHD2 patients do not have D4Z4 contraction but do also reveal the more relaxed chromatin structure and subsequent DUX4 expression. Recently, Lemmers et al. identified the causal determinant for most FSHD2 patients and have clearly demonstrated that mutations in the epigenetic modifier SCHMD1 gene are a causal determinant for most FSHD2 patients by derepressing DUX4 transcription (26). DUX4 mRNA expression is so far the only molecular determinant in common between FSHD1 and FSHD2 patients, which strongly suggests its direct implication in the FSHD phenotype.

Despite the fact that DUX4-FL is always expressed in FSHD myotubes, the consequence of this expression on muscle function is not yet clearly understood and the link between the DUX4 expression and the development of FSHD is not well established.

We had access to quadriceps and trapezius muscle biopsies

isolated from a 14-week FSHD1 fetus from which we established primary muscle cultures and investigated whether some molecular hallmarks of the pathology were already present in fetal biopsies and muscle cultures derived from these biopsies. We observed that DUX4-FL is expressed in the myogenic cells derived from the two FSHD1 biopsies and interestingly, myotubes derived from FSHD1 quadriceps express four times more DUX4-FL than myotubes derived from FSHD1 trapezius. Moreover, some genes downstream of DUX4-FL were specifically activated in FSHD1 biopsies and not in controls, thus showing that a molecular signature potentially leading to FSHD1 phenotype is already present in the developing fetal muscle. Finally, we observed that TRIM43 and MBD3L2, which are already overexpressed in FSHD1 fetal quadriceps muscle biopsies, are also expressed in adult quadriceps biopsies at similar levels to those observed in FSHD1 fetal quadriceps muscle biopsies, thus opening a new field of investigation concerning the molecular mechanisms leading to FSHD.

RESULTS

DUX4-FL is expressed in cells isolated from fetal FSHD1 muscles

As DUX4 has been described to be expressed in differentiated cells, we used myotubes at day 5 of differentiation and compared DUX4 mRNA expression in myoblasts derived from either fetal FSHD1 trapezius (FTrap) or quadriceps (FQuad) and quadriceps control cells (MCQ). Three primer sets were used (Fig. 1A): DUX4-all primers are able to amplify all the DUX4 mRNA isoforms described previously (16), DUX4-FL primers can only amplify the DUX4 long isoform where intron 1 is retained (Fig. 1B) and DUX4-UTR primers are flanking the two introns of DUX4 and can amplify both the DUX4-FL isoforms. When DUX4-all primers were used, we observed that the total expression of the DUX4 isoform mRNAs was 5-fold lower in FTrap as compared with FQuad myotubes (Fig. 1C). In the control myotubes (MCQ), DUX4 was also observed but its expression was 10-fold less than that in FQuad. When specific primers to DUX4-FL2 were used, we observed that the expression of DUX4-FL2 was 3-fold lower in the FTrap than in the FQuad (Fig. 1C) and DUX4-FL2 was never detected in control myotubes (Fig. 1C and data not shown, performed on three other control samples). Finally no amplification was observed in the control cells with the DUX4-UTR primers whereas the two DUX4-FL isoforms were observed in the FSHD fetal cells (Fig. 1D).

These results show that both the DUX4-FL isoforms are already expressed in muscle cells isolated from FSHD1 fetal biopsies. This result was confirmed at the protein level: an immunostaining against DUX4 was carried out on FQuad and revealed that 8.89% of the nuclei within myotubes were DUX4 positive (Fig. 1E).

FSHD fetal quadriceps and trapezius muscles have different differentiation kinetics during development

In order to determine whether this differential DUX4 mRNA expression could be owing to a difference in myotube formation,

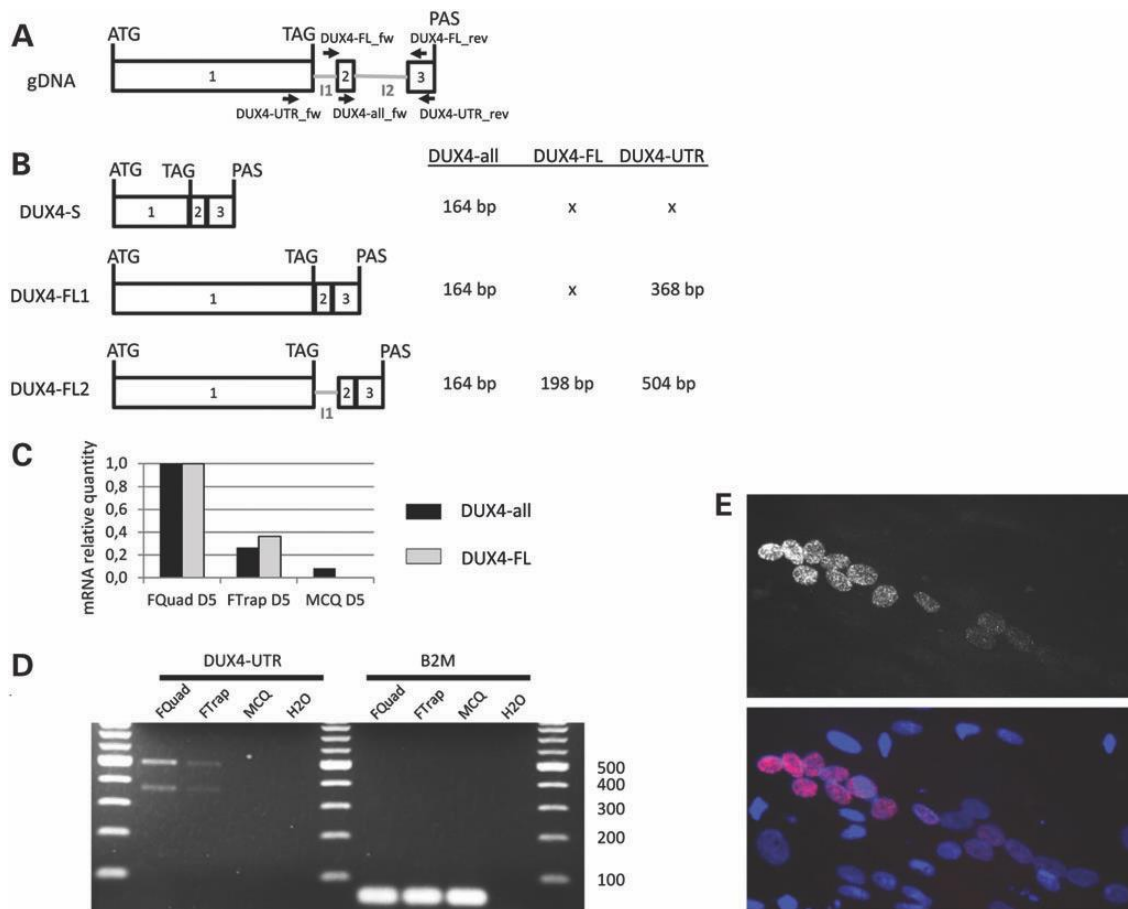


Figure 1. The expression of DUX4-FL mRNA in FSHD and control fetal myotubes. (A) Schematic representation of the DUX4 gene and the resulting DUX4 mRNA expression in muscle. The two primer locations that have been used are represented. (B) Representation of the DUX4 mRNA isoforms previously described (16). The primers for DUX4-all are able to amplify all of the DUX4 isoforms whereas the primers for DUX4-FL cannot amplify DUX4-s or DUX4-FL1. The PCR product sizes are indicated. (C) DUX4 relative expression in fetal-derived FSHD1 (FQuad and FTrap) and control (MCQ) myotubes after 5 days of differentiation. A RT-qPCR (taqman assay) was carried out on 1 mg of total RNA using either DUX4-all (black bars) or DUX4-FL (gray bars) primer sets. The results obtained with FQuad were used as a reference (and arbitrarily put at 1) and were normalized using B2M. (D) The amplification products using the DUX4-UTR or B2M primers were separated on a 2% agarose gel. The expected sizes of the PCR product were 368 bp when both introns 1 and 2 are spliced and 504 bp when intron 1 is retained. The molecular weight is a 100 bp ladder. The exons of DUX4 are indicated as gray boxes numbered 1 to 3. I: Intron; gDNA: genomic DNA. (E) Immunostaining against DUX4 was realized on FQuad 5 days after the induction of differentiation. On the upper panel, DUX4 was labeled with the E5.5 antibody (dilution 1/10). In the lower panel, nuclei have been counterstained with Dapi. The percentage of DUX4-positive nuclei represents 8.89% of the total number of nuclei within myotubes.

differentiation kinetics were analyzed. Cells were harvested at different time points after differentiation and DUX4-all or DUX4-FL2 mRNA were assessed by qPCR. For FQuad, DUX4-all and DUX4-FL2 mRNA expression increased during differentiation until day 4 and then decreased (Fig. 2A). For FTrap, the expression pattern of DUX4-all mRNA was delayed, with an increase until day 5–7 when a plateau was observed (Fig. 2B). As previously observed, the expression levels of DUX4-all and DUX4-FL were much higher (4–5-fold) in FQuad as compared with FTrap independently of the differentiation time point.

As we observed a delay in the expression of DUX4 mRNA in FTrap as compared with FQuad, we investigated whether the kinetics of differentiation/fusion were similar in both the cultures. The expression of Myogenin (MyoG), a transcriptional activator of the myogenic program that permits the transition from proliferating myoblasts to differentiating myotubes (for review see (27)), was compared between FQuad and FTrap

differentiated cultures. Interestingly, differences were observed: in FQuad cultures, the peak of expression of MyoG mRNA was observed at D1 of differentiation and then its expression decreased (Fig. 3A). In FTrap cultures, the MyoG peak was observed at D4, suggesting a delay in the activation of the genes required for differentiation and thus of the whole differentiation program. This delay in differentiation was confirmed using the inclusion of BIN1 exon11 in BIN1 mRNA. Inclusion of the exon 11 is specific of muscle differentiation, and this generates an isoform of BIN1 implicated in tubular invaginations of the membrane (28). Exon 11 inclusion was clearly observed at D1 for FQuad, whereas it was included only from D4 in FTrap (Fig. 3B and C). These results strongly suggest that there is a delay in the activation of the myogenic program in FTrap as compared with FQuad.

We next investigated whether the differences we observed between FQuad and FTrap are related to the pathology or to an intrinsic difference between the different muscle groups

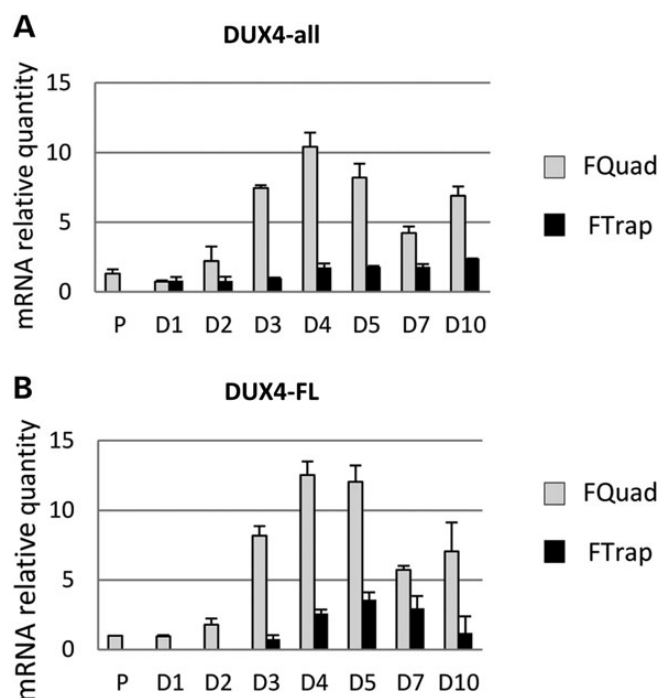


Figure 2. The expression of DUX4 mRNA in FTrap and FQuad during myotubes formation. Differentiation kinetics was performed on FQuad and FTrap cultures. Cells were harvested at different time points. Total RNA was extracted using trizol, and RT-qPCRs using either Dux4-all (A) or DUX4-FL (B) primers were performed. The results obtained with FQuad in proliferation were used as a reference and were normalized to B2M. The experiments were repeated in duplicate. P: proliferation; D: day after the induction of differentiation. Error bars represent SD of PCR triplicate.

(quadriceps versus trapezius). The same experiments were performed with cells isolated from fetal control biopsies (i.e. without any neuromuscular disease). We searched for DUX4 mRNA expression during the kinetics of differentiation by qPCR but never observed a consistent expression of DUX4-FL mRNA although sporadically amplification could be detected, thus confirming previous reports (17,29). As for the FSHD cultures, we also observed in four control cultures a delay in MyoG mRNA expression (Supplementary Material, Fig. S1), thus confirming that the delay we had initially observed in the FSHD1 fetal trapezius muscle is owing to an intrinsic difference in muscle maturation between trapezius and quadriceps rather than being directly linked to FSHD.

Genes downstream of DUX4-FL are activated in myotubes derived from FTrap and FQuad

DUX4-FL overexpression in vitro potentially induces the expression of several genes (14,24). As DUX4-FL mRNA expression levels are different in FTrap and FQuad myotubes, we investigated the expression of these DUX4-FL target genes in cultures derived from FTrap and FQuad. The expression of MBD3L2, ZSCAN4, TRIM43, DEFB103 and ZNF217 mRNAs was analyzed by qPCR. All of these genes, except ZNF217, were activated during differentiation (Fig. 4A). It should be noted that the kinetics of gene expression were different during differentiation of cultures derived from FTrap and

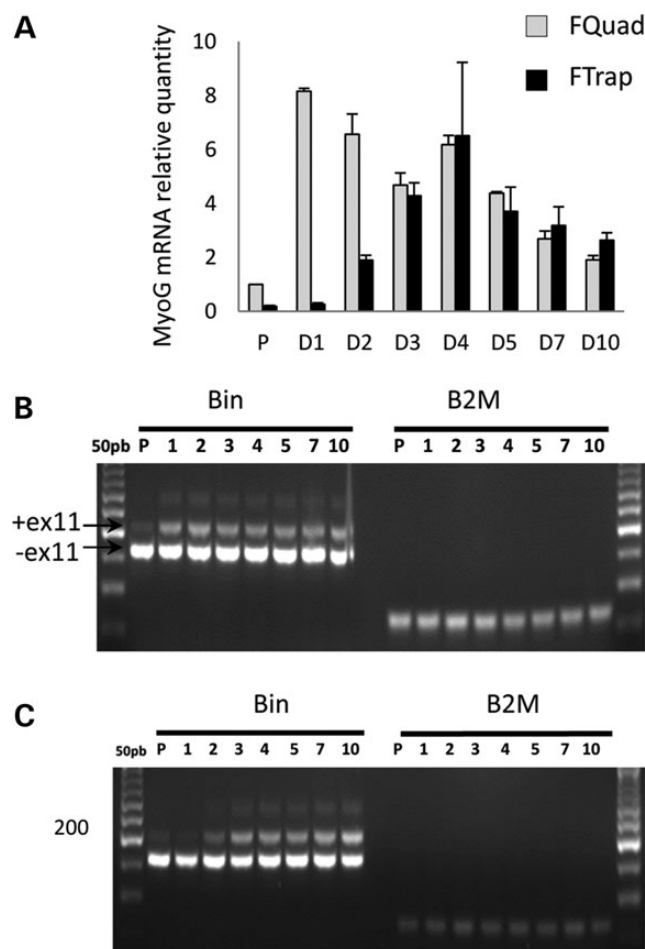


Figure 3. Characterization of the differentiation kinetics in FQuad and FTrap. To analyze differentiation kinetics on FQuad and FTrap, total RNA was extracted at different time points after the induction of differentiation by serum starvation. RT-qPCR was performed for the detection of MyoG (myogenin) (A) and BIN1 exon 11 (B FQuad; C FTrap). B2M was used as the normalizer.

FQuad: whereas the expression of target genes was maximal at day 4 of differentiation in FQuad cultures, this peak was delayed to day 5 to 7 in FTrap cultures, as observed for DUX4 mRNA expression. The relative expression levels of MBD3L2 and ZSCAN4 mRNA were next compared between FQuad and FTrap cultures (Fig. 4B). We observed that the levels of MBD3L2 and ZSCAN4 expression were 3–5-fold lower in FTrap than in FQuad cultures, as was observed for DUX4 mRNA expression.

FQuad and FTrap myoblasts were next transduced using a lentiviral vector containing an shRNA against all DUX4 mRNA isoforms. In cells transduced with the DUX4 shRNA (Fig. 5A), a down-regulation of DUX4-FL mRNA was observed (21 and 44% of residual mRNA in FQuad and FTrap myotubes, respectively, at day 5 of differentiation), demonstrating the efficacy of the shRNA. The depletion of DUX4 was also demonstrated at the protein level using a DUX4-specific antibody (Supplementary Material, Fig. S2). We next investigated the expression level of MBD3L2, ZSCAN4, DEFB103, TRIM43 and ZNF217 in the transduced cells (Fig. 5B). In the cells transduced with the empty shRNA, the mRNA expression level of the target genes

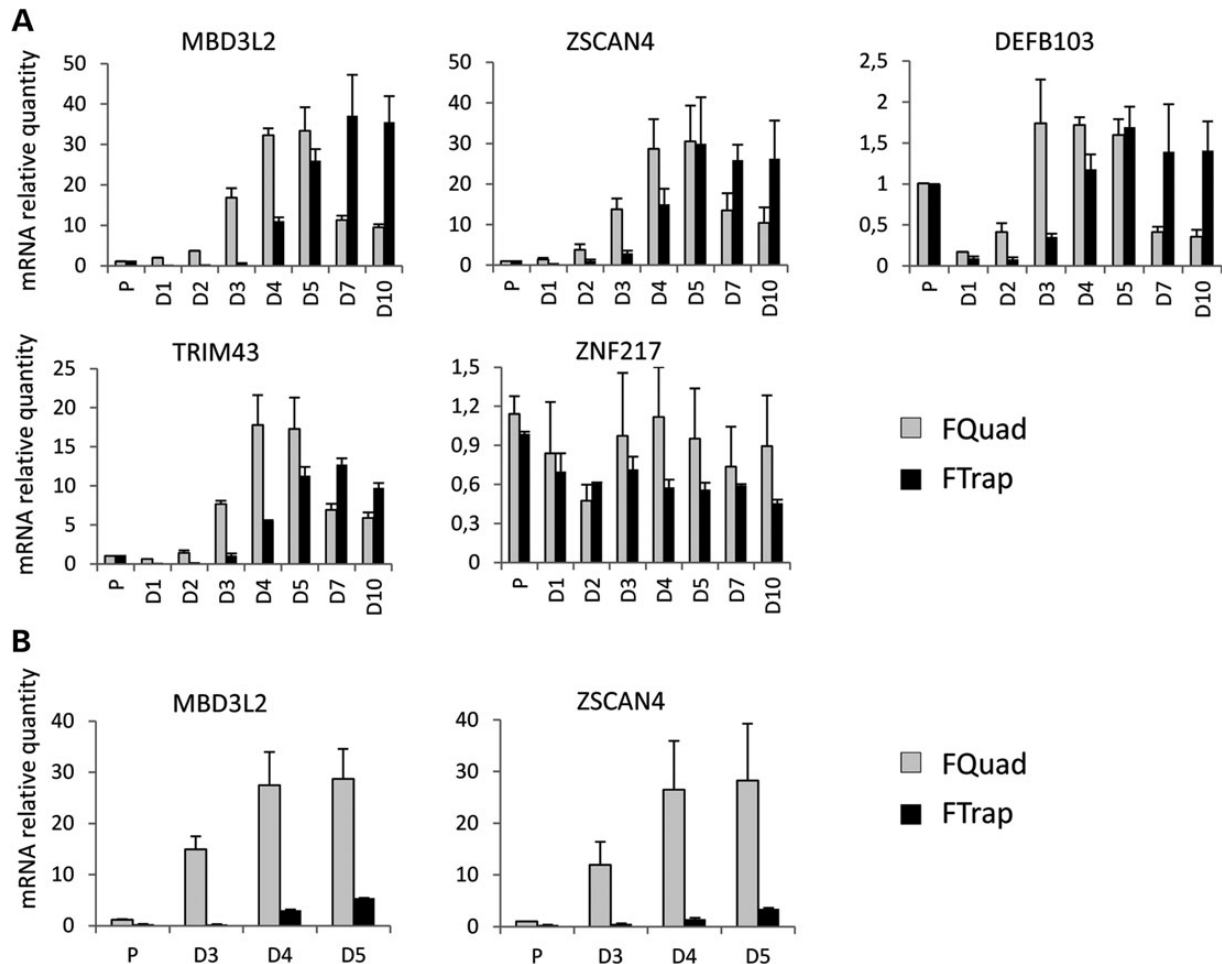


Figure 4. The expression level of genes downstream of DUX4 in FQuad and FTrap. Real-time RT-PCRs were carried out at different time points after the induction of differentiation. Several genes described as being activated by DUX4 were tested. The experiment was made on FQuad (black bars) and FTrap (gray bars). (A) Values measured in each cell type were normalized to its level in proliferation condition. (B) The results are normalized to the level of expression in proliferating FQuad. FTrap expression level is thus relative to FQuad, which allows a better comparison between both differentiated cultures. Error bars represent SD of PCR triplicates carried out on two independent kinetic experiments.

(except for DEFB013 in FQuad) increases during differentiation as expected. As previously observed in Figure 4A, the differentiation process did not modulate the expression level of ZNF217, thus suggesting that ZNF217 expression is not regulated by DUX4. This was confirmed in the cells transduced with the shDUX4, where the mRNA expression level of ZNF217 was not affected by the presence of the shRNA (Fig. 5B), whereas for the other genes, the presence of the DUX4 shRNA inhibited their overexpression during differentiation, thus confirming the down-regulation of the DUX4 expression that turns down its target genes. The levels of expression of DUX4-FL downstream targets further confirm the different levels of DUX4 expression observed in the differentiated cultures.

As many genes have previously been described to be modulated by the overexpression of DUX4-FL, we investigated the impact of the shDUX4 expression in FQuad and FTrap on the expression of TP53, THOC4, DBR1, TFI11, SRSF8, PRAMEF1 and MURF1 (Supplementary Material, Fig. S3). None of these genes were modulated except PRAMEF1, which was

dramatically down-regulated in both FQuad and FTrap when the shDUX4 is expressed and MURF1 mRNA, which was up-regulated.

Genes downstream of DUX4-FL are also specifically activated in FSHD1 fetal biopsies

Broucquault et al. (29) have shown previously that DUX-FL is expressed in fetal FSHD muscle biopsies and because we have used the same FSHD biopsies as them we decided to investigate the mRNA expression level of genes downstream of DUX4-FL. Interestingly, we found that MBD3L2 and TRIM43 mRNA expression levels were systematically higher in the fetal FSHD biopsies (18–36-fold and 10–26-fold for MBD3L2 and TRIM43, respectively), indicating that these DUX4-FL downstream genes are specifically activated in FSHD samples (Fig. 6). However, DEFB103, ZSCAN4, PRAMEF1 and ZNF217 mRNA expression levels were not different in FSHD biopsies as compared with control biopsies. These results

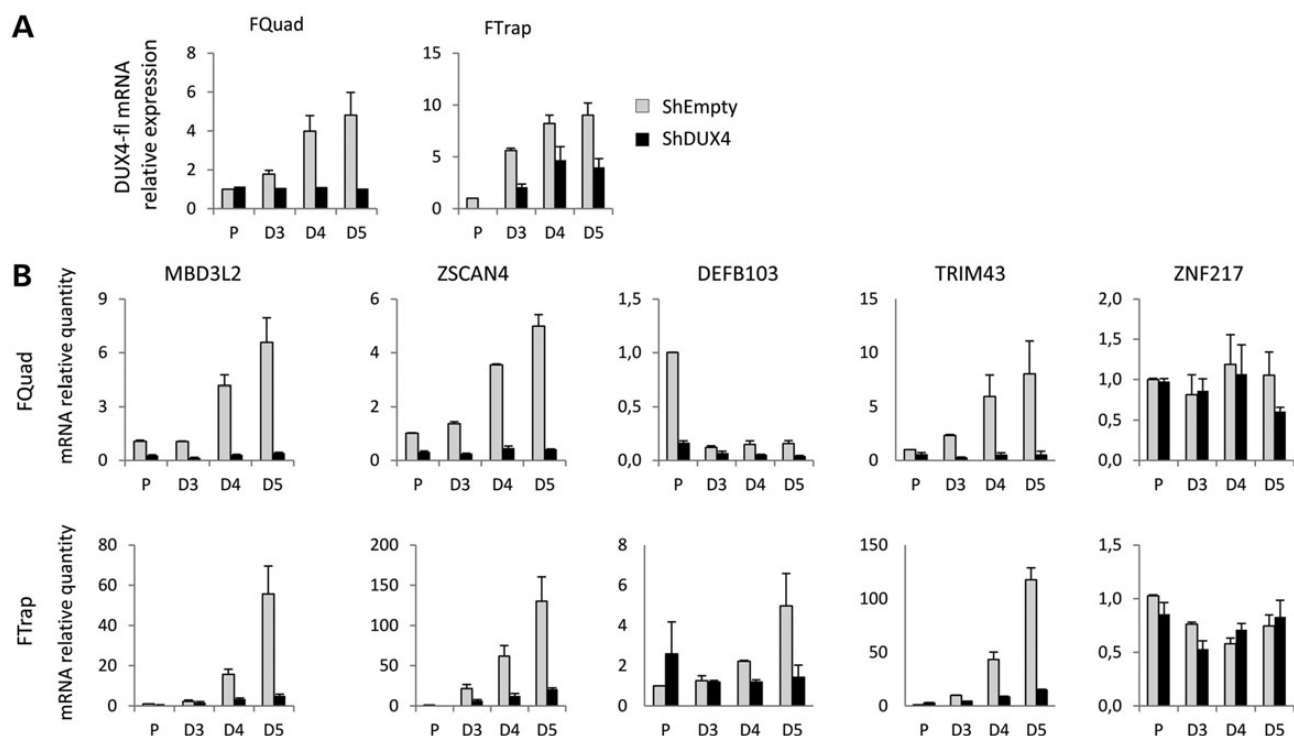


Figure 5. DUX4 inhibition using shDUX4 in FQuad and FTrap and activation of genes downstream of DUX4. A lentivirus (pLL3.7) carrying both an shRNA directed against DUX4 and a CMV-GFP cassette was used to transduce FQuad (left panel) or FTrap (right panel) cultures. After sorting of the CD56⁺ (a marker for myogenic cells) and GFP⁺ double positive cells, a differentiation kinetic was carried out. (A) DUX4 mRNA expression was analyzed by real-time PCR using the DUX4-all primers in the FQuad (left panel) and FTrap (right panel) transduced cells. At day 5 of differentiation, a student's t-test reveals that $P \leq 0.04$ and 0.03 for FQuad and FTrap, respectively. (B) The mRNA expression level of the downstream target genes was assessed by real-time PCR on FQuad (upper panel) and FTrap (lower panel) cultures transduced with either the shDUX4 (black bar) or the shEmpty (gray bar). A T-test realized at day 5 of differentiation reveals that $P \geq 0.05$ for MBD3L2, ZSCAN4, TRIM43 and DEFB103. Error bars represent SD of PCR triplicate realized on two independent experiments.

demonstrate that a modification in the expression of some but not all of the genes downstream of DUX4-FL is already present in fetal FSHD1 biopsies, strongly suggesting that DUX4-FL is also expressed during fetal development.

Expression of genes downstream of DUX4-FL is similar in fetal and adult samples

We next investigated the mRNA expression levels of MBD3L2, ZSCAN4, ZNF217, PRAMEF1, DEFB103 and TRIM43 in adult FSHD and control biopsies (Fig. 7A). We did not see any difference for ZNF217 and DEFB103. The overexpression of ZSCAN4 and PRAMEF1 was not confirmed in all adult samples. Finally, we did observe a clear modification in MBD3L2 and TRIM43 mRNA expression levels because the means of mRNA overexpression were 63 and 163, respectively. It should be noted that the FSHD1 samples are heterogeneous with sample #FSH1 always showing an important activation of the genes studied and sample #FSH3 always showing a very low activation of these genes. As FSHD1 has not been described to be a fetal pathology, we investigated whether the mRNA expression levels of TRIM43 and MBD3L2 were similar in adult and fetal biopsies. As shown in Figure 7B, we observed that the expression levels of both genes were comparable in fetal and adult FSHD biopsies.

DISCUSSION

Recent studies have suggested that DUX4 is a central player in FSHD pathophysiology. Indeed, the isoform DUX4-FL is expressed in myotubes isolated not only from FSHD1 but also from FSHD2 patients and has been shown to be toxic for myonuclei (14,16,20,26,30). Nevertheless, despite the fact that DUX4 might play an important role in the pathology, the mechanisms by which DUX4 is involved in FSHD still needs to be deciphered. One key to understanding its action is to clearly define when it first begins to be expressed in the skeletal muscle. Clinical features of FSHD usually appear in the second decade of life. However, a minority of patients carrying the most extended contraction develop the disease during late childhood. Nevertheless, nothing is known about the possible effect the D4Z4 contraction could have during fetal muscle growth and development.

In this article, we investigated the DUX4 expression during fetal development and show that DUX4-FL mRNA is already expressed in primary myogenic cells derived from two muscle biopsies derived from a single FSHD1 fetus (owing to the difficulty in obtaining FSHD fetal cells, our in vitro study was restricted to 1 fetus). Interestingly, the percentage of nuclei expressing the protein DUX4 does not seem to be different in fetal and in adult myotubes: DUX4 is present in almost 9% of the nuclei in fetal cells, and it was previously published that between 0.5 and 10% of the nuclei are DUX4 positive in adult

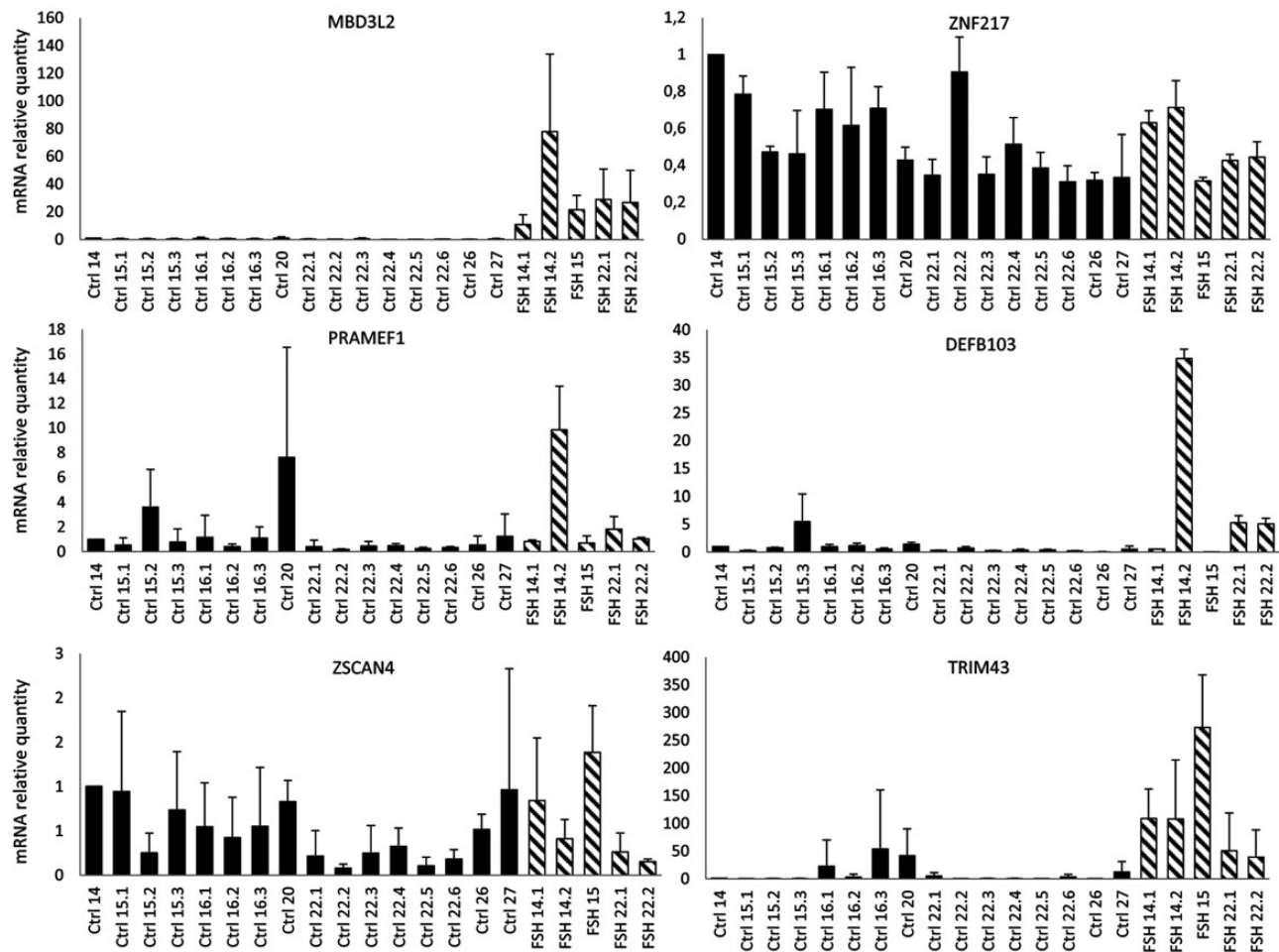


Figure 6. The expression level of genes downstream of DUX4 in quadriceps fetal biopsies. Quadriceps biopsies from five FSHD1 fetuses (hatched bars) and 16 quadriceps control biopsies obtained at different stages of development (back bars) were analyzed. A RT-q PCR was made on each mRNA preparation to determine their relative quantity. B2M was used as the normalizer gene. The result obtained with the Ctrl14 was used as the reference to generate arbitrary units in order to compare results between samples. Statistical analyses (T-test) revealed that the mean expression level for MBD3L2 and TRIM43 was different in control and FSHD1 samples ($P = 0.02$ and 0.03 , respectively).

cells (30,31). However, owing to the limited number of control cells available for this study, we cannot speculate that DUX4-FL is not expressed in fetal control cells as it was previously demonstrated with adult cells (17). Actually, Broucqsaute and colleagues (29) have observed a DUX4-FL mRNA expression in 8 of 27 fetal control biopsies. Several groups have previously shown that DUX4-FL mRNA is expressed in muscle cells derived from adult FSHD1 muscle biopsies (12,14,16, 17,26,32), but this is the first time that the DUX4 expression has been demonstrated in fetal FSHD1 cells. The two DUX4-FL isoforms, in which the first intron is spliced or not (DUX4-FL1 and DUX4-FL2, respectively), have been described to be expressed in both adult FSHD1 and FSHD2 myotubes, and we have confirmed this result on fetal FSHD1 myotubes. Interestingly, independent of the DUX4 primer set that was used, the DUX4 mRNA expression level was consistently lower in the trapezius than that in the quadriceps. This last result was surprising because the trapezius muscle is described to be one of the first affected muscles and so we were expecting to see a higher level of DUX4 mRNA in this muscle compared

with the quadriceps. Several possibilities could explain this result: (i) as FSHD is a disease affecting muscles asymmetrically by an unknown mechanism, it is possible that the biopsies were realized on a future poorly affected trapezius and a highly affected quadriceps and (ii) DUX4 expression is not the only determinant that explains the pattern of the affected and non-affected muscles in FSHD patients. This is in agreement with the fact that DUX4-FL has also been described to be expressed in control myotubes and biopsies (17,29). A third possibility is that the DUX4 expression is more toxic in Trapezius than that in Quadriceps owing to intrinsic difference between these two muscles. When a myonucleus starts to express DUX4, this will trigger apoptosis and cell death, thus decreasing the number of DUX4-positive nuclei and global DUX4 mRNA expression in FTrap, whereas FQuad myonuclei would not trigger apoptosis in response to the DUX4 expression.

Recently, numerous target genes downstream of DUX4 have been identified after transduction of control primary myoblasts by a lentiviral vector expressing DUX4-FL (24). Similarly, of the five genes we tested (MBD3L2, ZSCAN4, DEFB103,

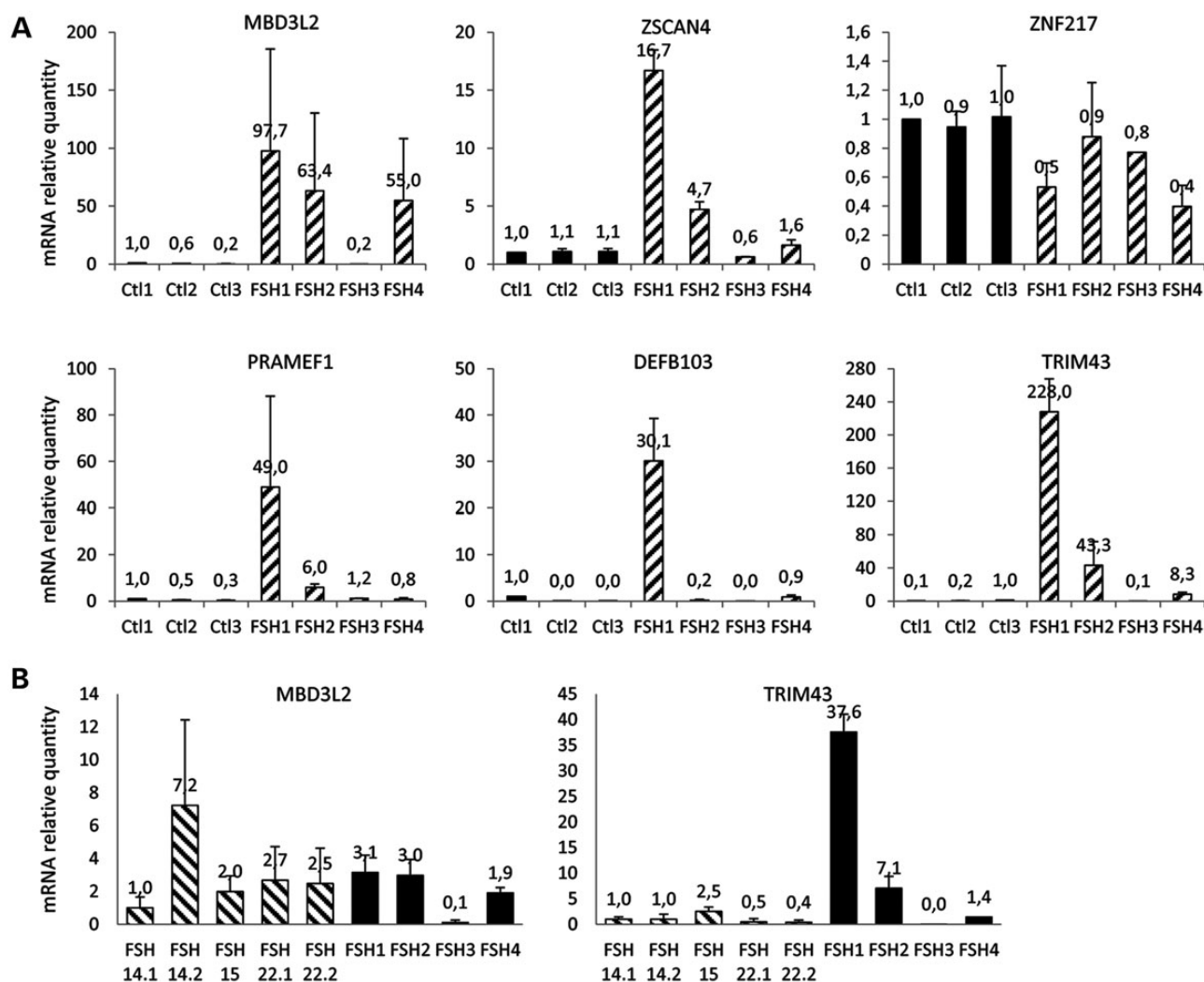


Figure 7. The expression level of genes downstream of DUX4 in quadriceps adult biopsies. (A) Four FSHD1 (hatched bars) and three healthy control (black bars) adult quadriceps biopsies were analyzed. For each gene, the result obtained with the Ctrl1 was used as the reference for arbitrary quantification. B2M gene was used as normalizer. Statistical analysis (Student's T-test) reveals that $P \geq 0.05$ only for MBD3L2 and TRIM43. (B) MBD3L2 and TRIM43 expression levels were compared in five fetal (hatched bars) and four adult (black bars) FSHD1 biopsies. Statistical analysis (t-test) does not show any difference in adult and fetal samples ($P \geq 0.22$ for MBD3L2 and $P \geq 0.16$ for TRIM43).

TRIM43 and ZNF217), all of them except ZNF217 were up-regulated in the fetal cultures FQuad and FTrap during differentiation. However, in fetal FSHD1 biopsies, we observed that only MBD3L2 and TRIM43 were specifically up-regulated. Considering that FSHD induces a muscle atrophy, which is often asymmetric, it is possible that the modification of the expression of the other genes (ZSCAN4, DEFB103 and PRAMEF1) in the FSHD fetal biopsies is influenced by this asymmetry. Moreover, in adult FSHD1 biopsies, we also confirmed that MBD3L2 and TRIM43 are up-regulated. These two genes can therefore be considered as robust markers for the DUX4 expression in both fetal and adult samples. It should also be noted that TRIM43 and MBD3L2 expression levels were similar in fetal and adult FSHD1 biopsies. The molecular events leading to the FSHD phenotype are thus already initiated during fetal development. Because it is difficult to imagine that all five FSHD fetuses would have been infantile cases, which represent 5–10% of all the FSHD cases, our data open a new field of

investigation for mechanisms leading to FSHD. Indeed, FSHD is described as an adult onset myopathy with clinical signs of the disease only appearing during the second decade of life. One hypothesis is that muscle fibers can only tolerate a given amount of DUX4 during a given amount of time, beyond which the delicate balance of muscle maintenance and repair will be lost. In this case, the amount of DUX4 expressed in each fiber is the same throughout life but results in fiber atrophy only after several years once the muscle is no longer able to compensate for the toxic effect of DUX4. FSHD could thus be at least partly attributed to the accumulation of DUX4 and genes downstream of DUX4 throughout life. An alternative hypothesis would result from DUX4 acting already by some epigenetic mechanism during fetal skeletal muscle development without having any major consequences on myogenesis and fiber growth, until a secondary event potentiates its toxic effect.

In conclusion, our results demonstrate that DUX4-FL is already expressed during fetal muscle development, as

evidenced in cultures derived from fetal biopsies, but more importantly by the activation of DUX4 target genes in these biopsies. Whether it is a priming event that requires additional signals or the accumulation of this signal over time that triggers the FSHD phenotype we do not know, and more research will be needed before we understand the complexity of FSHD pathophysiology.

MATERIALS AND METHODS

Ethic statement

This study and the collection of fetal and FSHD1 muscle biopsies was approved by the ‘Agence Française de la Biomedecine’ of the Ministry of Health to have legal access to the biological material in full accordance with the law (research protocol number PFS12–007). Therapeutic abortions are following a very precise protocol: after discussions with the medical team, the decision can be made to terminate the pregnancy. The parents can then choose to sign an informed consent for a fetal autopsy that will be carried out by pathologists. If the parents indicate that they accept that biological samples are taken for scientific investigations within the frame of scientific research on normal and pathological prenatal development, muscle biopsies will be taken. They will be used either to generate primary cultures of myoblasts or to be frozen for further molecular analyses.

FSHD patient reports

The FSHD1 fetuses were aborted for medical reasons. The fetal muscle biopsies of FSH14.1 fetus were performed by fetopathologists at AP-HP (Assistance Publique—Hôpitaux de Paris). The father was diagnosed for FSHD1 at age 23 and displayed a typical clinical phenotype including facial and scapula fixator muscle weakness. He carried four repeated units on 4qA chromosome 4. He was 32 when the fetus, carrying the same contraction, was aborted and at present (2 years later), the father can still walk but uses a cane. The genetic anomaly was inherited from his father who carried the 4QA D4Z4 contraction. The FSH14.2 fetus was aborted at 14 weeks of development and carried 1.5 D4Z4 repeats. The FSH15 was aborted at 15 weeks of development and carried four D4Z4 repeats. The fetuses FSH22.1 and 22.2 were aborted at 24 weeks of development and they both carried seven D4Z4 repeats.

Adult quadriceps muscle biopsies were obtained from typical FSHD1 patients and healthy control individuals using standardized muscle biopsy protocol. FSH1 and FSH2 were, respectively, 59- and 60-year-old men carrying seven D4Z4 repeats with a clinically non-affected quadriceps. FSH3 was a 62-year-old man carrying seven D4Z4 repeats with clinically affected quadriceps muscles. FSH4 was a 64-year-old woman carrying six D4Z4 repeats with mildly affected quadriceps. Healthy controls were selected in the same age and sex range as the FSHD patients.

Muscle biopsies and cells

Primary cells from FSH14.1 quadriceps (FQuad) and trapezius (FTrap) muscles were derived from the different muscle biopsies coming from the same fetus. Briefly, the biopsies were carefully minced, and the explants were plated in culture dishes previously coated with FBS. Once the cells had migrated from the explants

to colonize the culture dish, they were trypsinized and the percentage of CD56-positive cells was determined by FACS analysis. Briefly, 50 000 cells were incubated with 1 ml of CD56 antibody (CD56-APC, Becton Dickinson, Le Pont de Claix, France) in PBS FBS 2%. A FACSCalibur (Becton Dickinson, Le Pont de Claix, France) was used to determine the myogenicity, i.e. percentage of CD56-positive cells. If this percentage was lower than 80%, the myogenic cells were enriched in CD56-positive cells using MACS columns according to the manufacturer instructions (Miltenyi Biotec, Paris, France).

Cell culture

Fetal FSHD1 and control primary cells were cultivated in proliferation medium (4 vol of DMEM, 1 vol of 199 medium, FBS 20%, Gentamycin at 50 mg/ml, Life technologies, Saint Aubin, France). The differentiation was induced by replacing the proliferation medium by DMEM supplemented with insulin (10 mg/ml). The myogenicity, defined as the percentage of CD56-positive cells, was checked before any experiments performed with the cells, and the cells were always enriched in CD56+ cells using MACS columns immediately before plating if this percentage was \geq 80%.

Cell culture

After 5 days of differentiation, immunostaining was performed using the E5.5 antibody, which recognizes the C-terminus part of DUX4 (33). Cells were fixed with 4% paraformaldehyde for 10 min and permeabilized with 1% triton X-100 (Sigma) in PBS for 10 min at room temperature. Blocking was realized with PBS 2% FBS for 30 min at room temperature. DUX4 antibody (rabbit) was diluted 1/10 in PBS 2% FBS, and MF20 (Mouse IgG2B) was diluted 1/100. Cells were incubated overnight with DUX4 alone or both antibodies and washed three times with PBS 0.5% triton. Secondary antibodies were Ac GAM IgG2B-488 for anti-MF20 and GAR IgG H + L-555 (1/100) for DUX4. Incubation was carried out for 1 h in the dark. Finally, nuclei were counterstained with Hoescht (1/1000, Dako fluorescent mounting).

RNA extraction, PCR and real-time PCR

Total RNAs from either cells or human biopsies of muscles were extracted using trizol according to the manufacturer’s protocol (Life technologies, Saint Aubin, France). The quantity of RNA was determined using a nanodrop ND-1000 spectrophotometer (Thermo Scientific, Wilmington, DE, USA). The reverse transcription (RT) was carried out on 1 mg of total RNA using the Roche Transcriptor First Strand cDNA Synthesis Kit (Roche, Meylan, France) at 60°C for 50 min. In a final volume of 13 ml, 1 ml of oligo dT and 2 ml random hexameres were used as RT primers. All of the primers used in this study are listed in Table 1. With DUX4-UTR primers, the PCR was performed in a final volume of 25 ml with 1 ml of RT product, 1 ml of each forward and reverse primers at 20 pmol/ml. Thermal cycling conditions were 94°C for 5 min and then 36 cycles at 94°C for 30 s and 62°C for 45 s. The quantitative PCR was performed in triplicate with the probe mastermix (Roche, Meylan, France) on a LightCycler[®] 480 Real-Time PCR System (Roche, Meylan, France). DUX4 qPCR was performed in triplicate. Each well contained 0.25 ml of RT product, 4.5 ml of the probe mastermix 2X, 0.18 ml of each primers (20 pmol/ml) and

Table 1. Sequences of the primers

Targeted gene	Primer	Sequence	Size (bp)	Accession number
DUX4-all	DUX4-all_fw DUX4-all_rev	CCCAGGTACCAGCAGACC TCCAGGAGATGTAAGTCTAATCCA	164	HQ266761
DUX4-FL	DUX4-FL_fw DUX4-FL_rev	CCTGGGATTCCTGCCTTCTA AGCCAGAATTTACGGAAGA	198	HQ266761
DUX4-UTR	DUX4-UTR_fw DUX4-UTR_rev	AGGCGCAACCTCTCTAGAAAC TCCAGGAGATGTAAGTCTAATCCA	368 and 504	HQ266761
Myogenin	820U19 965L20	GAGTTCAGCGCCAACCCAG TGCCCGGCTTGAAGACAAT	165	NM_002479.5
BIN1	BIN F BIN R	AGAACCTCAATGATGTGCTGG CGTGGTTGACTCTGATCTCGG	162 and 207	NM_139351.2
MBD3L2	MBD3L2_fw MBD3L2_rev	CGTTCACCTCTTTTCCAAGC AGTCTCATGGGAGAGCAGA	142	NM_144614.3
ZSCAN4	ZSCAN4_fw ZSCAN4_rev	TGGGCCACTGCAATGACAA AGCTTCCTGTCCCTGCATGT	143	NM_152677.2
DEFB103	DEFB103_fw DEFB103_rev	GAGGATCCATTATCTTCTGTTTC AGGCAGCTGAGCACAGCA	135	NM_018661.3
TRIM 43	TRIM43_fw TRIM43_rev	ACCATCACTGGACTGGTGT CACATCCTCAAAGAGCCTGA	100	NM_138800.1
ZNF217	ZNF217_fw ZNF217_rev	AAGCCCTATGGTGGCTCC TTGATATGACACAGGCCTTTTC	99	NM_006526.2
PRAMEF1	PRAMEF1_fw PRAMEF1_rev	CTCCAAGGACGGTTAGTTGC AGTTCTCCAAGGGGTTCTGG	142	NM_023013.2
TP53	TP53_fw TP53_rev	GTTCCGAGAGCTGAATGAGG CTTGAGTCAGGCCCTTCTGT	159	NM_001126118.1
MURF1	MURF1_fw MURF1_rev	CTTGACTGCCAAGCAACTCA CAAAGCCCTGCTCTGTCTTC	80	NM_032588.3
THOC4	THOC4_fw THOC4_rev	CAGCAGACGTGCACTTTGAG TGACGTGACAAGCTGAATGTT	107	NM_005782.3
DBR1	DBR1_fw DBR1_rev	AAACCATGAAGCCTCAAATCA TTACCACACCAGCCAAACCT	92	NM_016216.3
SRSF8	SRSF8_fw SRSF8_rev	TTAGCTGGGCGCTGATTAGG AATGAAGCCAGGTCCCAGAG	71	NM_032102.2
TFIP11	TFIP11_fw TFIP11_rev	GCAGATGCCTTTCACAGGTT ACCATCGGGTCACAGTTCC	95	NM_012143.2

0.18 ml of the DUX4 probe (10 pmol/ml) in a final volume of 9 ml. qPCR cycling conditions were as follows: 95°C for 5 min, followed by 53 cycles at 95°C for 30 s and 62°C for 45 s. For the other genes, the qPCR was performed in a final volume of 9 ml with 0.4 ml of RT product, 0.18 ml of each forward and reverse primers 20 pmol/ml, and 4.5 ml of syber-green mastermix 2X (Roche, Meylan, France). The qPCR cycling conditions were 94°C for 5 min, followed by 50 cycles at 95°C for 30 s and 60°C for 15 s and 72°C for 15 s.

To amplify BIN1, PCR was performed using 2X ReddyMix PCR master mix (Thermo scientific, Rockford, IL, USA) with 1 ml of RT product, 0.5 ml of 20pmol/ml forward and reverse primers, in a final volume of 25 ml. The cycling conditions were as follows: 95°C for 5 min, followed by 32 cycles at 95°C for 20 s, 60°C for 20 s and 72°C for 20 s, then 72°C for 7 min, on a thermocycler GeneAmp PCR System 2700 (Applied Biosystems, Saint Aubin, France).

DUX4 shRNA construct and transduction

The sequence of DUX4 was analyzed for siRNA target using the Clontech software (<http://bioinfo.clontech.com/rnaidesigner/sirnaSequenceDesign.do>). One siRNA was selected and synthesized to be directly cloned into pLL3.7 (addgene, Plasmid 11795) using the hpaI and XhoI restriction enzyme sites. The oligonucleotides were TGGCAAACCTGGATTAGAGTTTCAA

GAGAACTCTAATCCAGGTTTGCCTTTTTTC for the sense strand and TCGAGAAAAAGGCAAACCTGGATTAGAGTTCTCTTGAACTCTAATCCAGGTTTGCCA for the anti-sense strand. The pLL3.7 shDUX4 vector was produced in human embryonic kidney 293 cells by quadri-transfection of plasmids encoding gag-pol proteins, Rev protein, envelop proteins (VSVg) and the transgene using PEI. 48 and 72 h later; the viral vector is filtered (0.22 µm) before being directly used to transduce myoblasts.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at HMG online.

ACKNOWLEDGEMENTS

We thank all the patients who provided the biopsies. We thank the Myobank and in particular Stephan Vasseur et Maud Chapart for providing the control adult biopsies. This work was supported by the Association Française contre les Myopathies (AFM-Téléthon, France), Université Pierre et Marie Curie (Emergence 2010 to J.D.), FSHD Global research foundation Ltd (to M.F.), the FSH society (FSHS-22012-03 to V.M.) and the Agence Nationale de la Recherche (FSHDecrypt,

ANR-09-GENO-038 to S.R.). The E5.5 antibody was a very kind gift from Dr S. Tapscott.

Conflict of Interest statement. None declared.

REFERENCES

1. Tawil, R. and Van Der Maarel, S.M. (2006) Facioscapulohumeral muscular dystrophy. *Muscle Nerve*, 34, 1–15.
2. van der Maarel, S.M., Frants, R.R. and Padberg, G.W. (2007) Facioscapulohumeral muscular dystrophy. *Biochim. Biophys. Acta.*, 1772, 186–194.
3. van Deutekom, J.C., Wijmenga, C., van Tienhoven, E.A., Gruter, A.M., Hewitt, J.E., Padberg, G.W., van Ommen, G.J., Hofker, M.H. and Frants, R.R. (1993) FSHD associated DNA rearrangements are due to deletions of integral copies of a 3.2 kb tandemly repeated unit. *Hum. Mol. Genet.*, 2, 2037–2042.
4. van der Maarel, S.M. and Frants, R.R. (2005) The D4Z4 repeat-mediated pathogenesis of facioscapulohumeral muscular dystrophy. *Am. J. Hum. Genet.*, 76, 375–386.
5. de Greef, J.C., Lemmers, R.J., Camano, P., Day, J.W., Sacconi, S., Dunand, M., van Engelen, B.G., Kiuru-Enari, S., Padberg, G.W., Rosa, A.L. et al. (2010) Clinical features of facioscapulohumeral muscular dystrophy 2. *Neurology*, 75, 1548–1554.
6. van der Maarel, S.M., Tawil, R. and Tapscott, S.J. (2011) Facioscapulohumeral muscular dystrophy and DUX4: breaking the silence. *Trends. Mol. Med.*, 17, 252–258.
7. Thomas, N.S., Wiseman, K., Spurlock, G., MacDonald, M., Ustek, D. and Upadhyaya, M. (2007) A large patient study confirming that facioscapulohumeral muscular dystrophy (FSHD) disease expression is almost exclusively associated with an FSHD locus located on a 4qA-defined 4qter subtelomere. *J. Med. Genet.*, 44, 215–218.
8. Scionti, I., Fabbri, G., Fiorillo, C., Ricci, G., Greco, F., D’Amico, R., Termanini, A., Vercelli, L., Tomelleri, G., Cao, M. et al. (2012) Facioscapulohumeral muscular dystrophy: new insights from compound heterozygotes and implication for prenatal genetic counselling. *J. Med. Genet.*, 49, 171–178.
9. Lemmers, R.J., Van Overveld, P.G., Sandkuijl, L.A., Vrieling, H., Padberg, G.W., Frants, R.R. and van der Maarel, S.M. (2004) Mechanism and timing of mitotic rearrangements in the subtelomeric D4Z4 repeat involved in facioscapulohumeral muscular dystrophy. *Am. J. Hum. Genet.*, 75, 44–53.
10. Lemmers, R.J., Wohlgemuth, M., Frants, R.R., Padberg, G.W., Morava, E. and van der Maarel, S.M. (2004) Contractions of D4Z4 on 4qB subtelomeres do not cause facioscapulohumeral muscular dystrophy. *Am. J. Hum. Genet.*, 75, 1124–1130.
11. Lemmers, R.J., de Kievit, P., Sandkuijl, L., Padberg, G.W., van Ommen, G.J., Frants, R.R. and van der Maarel, S.M. (2002) Facioscapulohumeral muscular dystrophy is uniquely associated with one of the two variants of the 4q subtelomere. *Nat. Genet.*, 32, 235–236.
12. Lemmers, R.J., van der Vliet, P.J., Klooster, R., Sacconi, S., Camano, P., Dauwerse, J.G., Snider, L., Straasheijm, K.R., van Ommen, G.J., Padberg, G.W. et al. (2010) A unifying genetic model for facioscapulohumeral muscular dystrophy. *Science*, 329, 1650–1653.
13. Gabriels, J., Beckers, M.C., Ding, H., De Vriese, A., Plaisance, S., van der Maarel, S.M., Padberg, G.W., Frants, R.R., Hewitt, J.E., Collen, D. et al. (1999) Nucleotide sequence of the partially deleted D4Z4 locus in a patient with FSHD identifies a putative gene within each 3.3 kb element. *Gene*, 236, 25–32.
14. Dixit, M., Ansseau, E., Tassin, A., Winokur, S., Shi, R., Qian, H., Sauvage, S., Matteotti, C., van Acker, A.M., Leo, O. et al. (2007) DUX4, a candidate gene of facioscapulohumeral muscular dystrophy, encodes a transcriptional activator of PITX1. *Proc. Natl. Acad. Sci.*, 104, 18157–18162.
15. Snider, L., Asawachaicharn, A., Tyler, A.E., Geng, L.N., Petek, L.M., Maves, L., Miller, D.G., Lemmers, R.J., Winokur, S.T., Tawil, R. et al. (2009) RNA transcripts, miRNA-sized fragments and proteins produced from D4Z4 units: new candidates for the pathophysiology of facioscapulohumeral dystrophy. *Hum. Mol. Genet.*, 18, 2414–2430.

16. Snider, L., Geng, L.N., Lemmers, R.J., Kyba, M., Ware, C.B., Nelson, A.M., Tawil, R., Filippova, G.N., van der Maarel, S.M., Tapscott, S.J. et al. (2010) Facioscapulohumeral dystrophy: incomplete suppression of a retrotransposed gene. *PLoS genetics*, 6, e1001181.
17. Jones, T.I., Chen, J.C., Rahimov, F., Homma, S., Arashiro, P., Beermann, M.L., King, O.D., Miller, J.B., Kunkel, L.M., Emerson, C.P. Jr et al. (2012) Facioscapulohumeral muscular dystrophy family studies of DUX4 expression: evidence for disease modifiers and a quantitative model of pathogenesis. *Hum. Mol. Genet.*, 21, 4419–4430.
18. Tupler, R., Berardinelli, A., Barbierato, L., Frants, R., Hewitt, J.E., Lanzi, G., Maraschio, P. and Tiepolo, L. (1996) Monosomy of distal 4q does not cause facioscapulohumeral muscular dystrophy. *J. Med. Genet.*, 33, 366–370.
19. Bosnakovski, D., Xu, Z., Gang, E.J., Galindo, C.L., Liu, M., Simsek, T., Garner, H.R., Agha-Mohammadi, S., Tassin, A., Coppee, F. et al. (2008) An isogenetic myoblast expression screen identifies DUX4-mediated FSHD-associated molecular pathologies. *Embo. J.*, 27, 2766–2779.
20. Vanderplanck, C., Anseau, E., Charron, S., Stricwant, N., Tassin, A., Laoudj-Chenivresse, D., Wilton, S.D., Coppee, F. and Belayew, A. (2011) The FSHD atrophic myotube phenotype is caused by DUX4 expression. *PloS one*, 6, e26820.
21. Wuebbles, R.D., Long, S.W., Hanel, M.L. and Jones, P.L. (2010) Testing the effects of FSHD candidate gene expression in vertebrate muscle development. *Int. J. Clin. Exp. Pathol.*, 3, 386–400.
22. Mitsuhashi, H., Mitsuhashi, S., Lynn-Jones, T., Kawahara, G. and Kunkel, L.M. (2013) Expression of DUX4 in zebrafish development recapitulates facioscapulohumeral muscular dystrophy. *Hum. Mol. Genet.*, 22, 568–577.
23. Kowaljow, V., Marcowycz, A., Anseau, E., Conde, C.B., Sauvage, S., Matteotti, C., Arias, C., Corona, E.D., Nunez, N.G., Leo, O. et al. (2007) The DUX4 gene at the FSHD1A locus encodes a pro-apoptotic protein. *Neuromuscul. Disord.*, 17, 611–623.
24. Geng, L.N., Yao, Z., Snider, L., Fong, A.P., Cech, J.N., Young, J.M., van der Maarel, S.M., Ruzzo, W.L., Gentleman, R.C., Tawil, R. et al. (2012) DUX4 activates germline genes, retroelements, and immune mediators: implications for facioscapulohumeral dystrophy. *Dev. Cell.*, 22, 38–51.
25. Wallace, L.M., Garwick, S.E., Mei, W., Belayew, A., Coppee, F., Ladner, K.J., Guttridge, D., Yang, J. and Harper, S.Q. (2010) DUX4, a candidate gene for facioscapulohumeral muscular dystrophy, causes p53-dependent myopathy in vivo. *Ann. Neurol.*, 69, 540–552.
26. Lemmers, R.J., Tawil, R., Petek, L.M., Balog, J., Block, G.J., Santen, G.W., Amell, A.M., van der Vliet, P.J., Almomani, R., Straasheijm, K.R. et al. (2012) Digenic inheritance of an SMCHD1 mutation and an FSHD-permissive D4Z4 allele causes facioscapulohumeral muscular dystrophy type 2. *Nat. Genet.*, 44, 1370–1374.
27. Faralli, H. and Dilworth, F.J. (2012) Turning on myogenin in muscle: a paradigm for understanding mechanisms of tissue-specific gene expression. *Comp. Funct. Genomics.*, 2012, 836374.
28. Lee, E., Marcucci, M., Daniell, L., Pypaert, M., Weisz, O.A., Ochoa, G.C., Farsad, K., Wenk, M.R. and De Camilli, P. (2002) Amphiphysin 2 (Bin1) and T-tubule biogenesis in muscle. *Science*, 297, 1193–1196.
29. Broucqsaault, N., Morere, J., Gaillard, M.C., Dumonceaux, J., Torrents, J., Salort-Campana, E., Maues De Paula, A., Bartoli, M., Fernandez, C., Chesnais, A.L. et al. (2013) Dysregulation of 4q35- and muscle-specific genes in fetuses with a short D4Z4 array linked to facio-scapulo-humeral dystrophy. *Hum. Mol. Genet.*, .
30. Tassin, A., Laoudj-Chenivresse, D., Vanderplanck, C., Barro, M., Charron, S., Anseau, E., Chen, Y.W., Mercier, J., Coppee, F. and Belayew, A. (2012) DUX4 expression in FSHD muscle cells: how could such a rare protein cause a myopathy? *J. Cell. Mol. Med.*, 17, 76–89.
31. Block, G.J., Narayanan, D., Amell, A.M., Petek, L.M., Davidson, K.C., Bird, T.D., Tawil, R., Moon, R.T. and Miller, D.G. (2013) Wnt/beta-catenin signaling suppresses DUX4 expression and prevents apoptosis of FSHD muscle cells. *Hum. Mol. Genet.*, .
32. Tsumagari, K., Chang, S.C., Lacey, M., Baribault, C., Chittur, S.V., Sowden, J., Tawil, R., Crawford, G.E. and Ehrlich, M. (2011) Gene expression during normal and FSHD myogenesis. *BMC. Med. Genomics.*, 4, 67.
33. Geng, L.N., Tyler, A.E. and Tapscott, S.J. (2011) Immunodetection of human double homeobox 4. *Hybridoma (Larchmt)*, 30, 125–130.

PUBLICATION N°2: Nuclear protein spreading: implication for pathophysiology of neuromuscular diseases

Ferreboeuf M., Mariot V, Furling D, Butler-Browne G, Mouly V, Dumonceaux J. Hum Mol Genet. 2014

Résumé

L'expression forcée de DUX4, l'un des gènes candidats pour la FSHD, a mis en évidence sa capacité à induire des dérégulations géniques majeures. Néanmoins, dans des conditions d'expression endogène, le niveau de DUX4 est extrêmement faible et est détecté dans environ 0,5 à 10 % des noyaux. Pourtant, le transcrit est lui détecté dans un nombre de noyaux beaucoup plus faible (1/1000) suggérant que la DUX4 est transcrite dans un faible nombre de noyaux mais que la protéine peut diffuser d'un noyau à l'autre.

Afin de tester cette hypothèse nous avons réalisé des essais de co-cultures entre des myoblastes Humain Contrôle ou FSHD et des myoblastes Murin (absence de DUX4). Nous avons ainsi mis en évidence par RNA-FISH et immunofluorescence que, bien que la transcription de DUX4 s'effectue dans un nombre restreint de noyaux Humain, la protéine est capable de diffuser entre les noyaux au sein des myotubes chimères générés (myotubes contenant à la fois des noyaux Humain et Murin). De plus, nous avons montré que la diffusion d'une protéine humaine dans des noyaux murins était capable d'activer la transcription de plusieurs gènes décrits comme dérégulés par DUX4 chez un modèle murin.

Ce modèle a ensuite été validé dans une autre pathologie, la dystrophie myotonique de type 1. D'un point de vue moléculaire, cette pathologie se caractérise par l'expansion de répétition CTG dans la 3'UTR du gène de la DMPK. Ces expansions titrent des facteurs d'épissage comme MBNL1, et forment des agrégats dans le noyau. La titration de MBNL1 entraîne des défauts d'épissage d'un grand nombre de gènes comme CLCN1 ou BIN 1 tout deux associés à la myotonie et la faiblesse musculaire chez les patients. Dans notre modèle de co-culture, l'ARNm DMPK muté (Humain) est capable de titrer des facteurs d'épissage murin et d'induire alors des défauts d'épissage de gènes Murin comme *Sercaf* ou *Bin1*.

Nos travaux ont ainsi démontré que, dans la FSHD, malgré le faible nombre de noyaux exprimant DUX4, la protéine est capable de diffuser entre les noyaux et d'y induire des dérégulations géniques assurant la dissémination du phénotype moléculaire pathologique dans un muscle entier.

Nuclear protein spreading: implication for pathophysiology of neuromuscular diseases

Maxime Ferreboeuf^{1,2,3,4,†}, Virginie Mariot^{1,2,3,4,†}, Denis Furling^{1,2,3,4}, Gillian Butler-Browne^{1,2,3,4}, Vincent Mouly^{1,2,3,4} and Julie Dumonceaux^{1,2,3,4,*}

¹Sorbonne Universités, UPMC Univ Paris 06, Myology Research Center UM76, F-75013 Paris, France, ²INSERM UMRS974, F-75013 Paris, France, ³CNRS FRE 3617, F-75013 Paris, France and ⁴Institut de Myologie, F-75013 Paris, France

Received March 13, 2014; Revised and Accepted March 18, 2014

While transfer of a protein encoded by a single nucleus to nearby nuclei in multinucleated cells has been known for almost 25 years, the biological consequences for gain-of-function diseases have not been considered. Here, we have investigated nuclear protein spreading and its potential consequences in two of the three most prevalent neuromuscular diseases. By performing co-cultures between diseased or control human myoblasts and murine C2C12 myoblasts, we demonstrate that in facioscapulohumeral dystrophy, although the transcription of the toxic protein DUX4 occurs in only a limited number of nuclei, the resulting protein diffuses into nearby nuclei within the myotubes, thus spreading aberrant gene expression. In myotonic dystrophy type 1, we observed that in human-mouse heterokaryons, the expression of a mutated DMPK from human nuclei titrates splicing factors produced by neighboring nuclei, inducing the mis-splicing of several pre-mRNAs in murine nuclei. In both cases, the spreading of the pathological phenotypes from one nucleus to another is observed, highlighting an additional mechanism that contributes to the dissemination and worsening of the muscle pathogenesis. These results indicate that nuclear protein spreading may be an important component of pathophysiology of gain of function muscular diseases which should be taken into consideration in the design of new therapeutic approaches.

INTRODUCTION

Therapies for neuromuscular diseases are usually based on the rescue of a diseased multinucleated muscle fiber by healthy nuclei or on the correction of diseased nuclei by gene therapy. These approaches are successful for loss-of-function diseases and are mediated by the spreading of normal proteins within the cytoplasm (1). However, such approaches may not be sufficient for gain-of-function diseases since the muscle fiber is a syncytium in which a nuclear protein encoded by a single nucleus can be transferred to nearby nuclei within the same fiber (2,3), thus spreading a phenotype.

In this study, we examined the effects of nuclear spreading on two gain-of-function muscular diseases. Facioscapulohumeral dystrophy (FSHD) and myotonic dystrophy (DM1) are two of the three most prevalent muscular diseases with an estimated prevalence of 4/100 000 and 4.5/100 000, respectively (<http://www.orpha.net>).

In FSHD, two loci have been described as being potentially involved in the pathological mechanism, leading to two forms of the disease, FSHD1 and FSHD2. The first one (FSHD1) is located on the subtelomeric part of chromosome 4 and corresponds to the contraction of a 3.3 kb repeated region called D4Z4, whereas control individuals carry between 11 and 150 D4Z4 repeats, FSHD1 patients carry between 1 and 10 D4Z4 repeats (4). The second one is located on chromosome 18 (FSHD2) where mutations in the SMCHD1 gene have been found, leading to epigenetic deregulations (5). These two loci are usually mutually exclusive but a few patients carrying both loci have been already characterized (6). Both loci induce chromatin relaxation allowing the expression of the transcription factor DUX4 [for review see (7)]. Indeed, each D4Z4 repeat contains the open reading frame (ORF) for the transcription factor DUX4 (8) and DUX4 thus appears to be the common point between FSHD1 and FSHD2 patients. However, there is some potential controversy regarding the expression of

DUX4, whereas DUX4 mRNA is only expressed in 1/1000 nuclei (9), DUX4 protein can be found in 0.5–9% of nuclei (10,11), suggesting that although DUX4 is not transcribed in most nuclei, the resulting protein may spread further. The progressive decrease in the signal intensity in consecutive nuclei systematically observed after DUX4 antibody staining would support this hypothesis (11,12), but such a spreading has never been demonstrated in FSHD. In this study, using RNA-FISH, we demonstrate that Dux4 mRNA is produced in only a very limited number of nuclei within myotubes. However, in co-cultures of human FSHD cells with murine C2C12, we show that the resulting DUX4 protein is present in the murine nuclei (which do not produce endogenous DUX4). In addition the human DUX4 activates murine genes downstream of DUX4, thus provoking a spreading of the pathological phenotype from diseased nuclei to neighboring healthy nuclei by nuclear protein transfer.

Consequences of nuclear spreading were also investigated in DM1, characterized by a repeat expansion in the DMPK gene localized on chromosome 19. The 3' untranslated region of this gene normally contains 5–37 copies of a CTG trinucleotide repeat while in DM1 patients, expansion of this unstable motif to 50–5000 copies have been reported leading to the expression of mutant CUG exp-RNAs that remain trapped in the nucleus as aggregates and sequester splicing factors such as MBLN1 [for review see (13)]. MBLN1 loss-of-function results in the misregulation of the alternative splicing of several pre-mRNAs including CLCN1 and BIN1, associated, respectively, with myotonia and muscle weakness (14,15). Using heterokaryons generated between human and murine myoblasts, we demonstrate that mutated DMPK mRNAs transcribed in the diseased nuclei titrate splicing factors in murine nuclei (which do not express the mutated RNA), thus inducing mis-splicing of several murine pre-mRNAs including MBLN1, BIN1 and SERCA1 in neighboring nuclei.

In both FSHD and DM1, we demonstrated the diffusion of the molecular pathological phenotype through nuclear protein spreading, highlighting an additional mechanism that contributes to dissemination and worsening of muscle pathogenesis. Nuclear protein spreading may be an important component of pathophysiology of gain-of-function muscular diseases which should be taken in consideration in therapeutic approaches.

RESULTS

DUX4-FL mRNA is expressed in only few nuclei within myotubes

To determine whether DUX4 transcription occurs in only a small percentage of myonuclei in our cellular model, RNA-FISH directed against DUX4 was performed on FSHD myotubes at Day 5 of differentiation. A probe generated by DIG-PCR allowing the recognition of both DUX4-FL1 (spliced intron 1) and -FL2 (non-spliced intron 1) isoforms was used (Fig. 1A). A very discrete number of myonuclei, with often only one nucleus per myotube, were found to express DUX4 (Fig. 1B and Supplementary Material, Fig. S1). In contrast using the DUX4 E5.5 antibody (16), we always observed several DUX4 protein-positive nuclei within each myotube. This demonstrates that DUX4 is transcribed in one or only very few nuclei within myotubes *in vitro*, but confirms that the resulting DUX4 protein spreads to the neighboring nuclei.

DUX4 spreads within myotubes

To verify if DUX4 is able to spread to several nuclei within a myotube, co-cultures between murine C2C12 cells (which do not express DUX4) and either FSHD or control human muscle cell lines were performed. At Day 5 of differentiation, cells were fixed and immunostainings were performed. Human nuclei were characterized by a bright human-specific LAMIN A/C labeling, whereas murine nuclei were identified by the presence of heterochromatin revealed after Hoechst treatment (Fig. 2A and B). Figure 2A confirms that murine and human myoblasts can fuse together *in vitro*, giving rise to myotubes containing both murine and human myonuclei. The presence of DUX4 was assessed by specific immunostaining (E5.5 antibody) as a control, no DUX4-positive nucleus was ever detected in co-cultures of murine C2C12 and human control myoblasts, which is in agreement with the fact that DUX4 is not expressed by either control human cells or murine cells (data not shown). In contrast, DUX4 protein was detected in heterokaryons formed in co-cultures of FSHD and C2C12 cells. DUX4 was seen in both human FSHD and murine nuclei and the consecutive gradient of DUX4 intensity often observed (Fig. 2B) indicates that DUX4 protein is able to spread from human into murine nuclei. The longest distance measured between human and murine DUX4-positive nuclei was 51 mm (Supplementary Material, Fig. S2).

DUX4 induces a diffusion of the pathological phenotype in myotubes

We next investigated whether the human DUX4 proteins were functional and able to activate genes downstream of DUX4 not only in human but also in murine nuclei. RT-PCRs were performed on mRNA extracted from the different co-cultures at Day 5 of differentiation. DUX4-FL2 was only detected in the presence of FSHD cells co-cultures (Fig. 3A) as expected. The expression of human genes downstream of DUX4 was analyzed by RT-qPCR. A 12.3- and 6-fold increase was observed for both human ZSCAN4 and DEFB103 [two genes previously shown to be activated by DUX4 in FSHD myotubes (17)] in co-cultures with immortalized FSHD cells when compared with co-cultures with control cells (Fig. 3B), showing that DUX4 is active in human nuclei. Interestingly, when the expression levels of ZSCAN4 and DEFB103 were analyzed in control or FSHD myotubes in the absence of murine cells, a stronger activation of the two genes was observed when compared with what was obtained in the presence of murine cells, thus suggesting that the diffusion of the pathological molecular phenotype is correlated to the number of nuclei being exposed to DUX4 (Supplementary Material, Fig. S3A). We also analyzed the expression level of two murine genes previously described to be DUX4 targets in a transgenic mouse over-expressing DUX4 (18). We confirmed after transfection of a DUX4-expression plasmid that Zscan4c and Wfdc3 are activated in the presence of DUX4 (Supplementary Material, Fig. S3). Using murine-specific primers (Fig. 3C), we detected a fold increase of 5 and 10.2 for murine Wfdc3 and Zscan4c, respectively, in C2C12-FSHD cell co-cultures when compared with C2C12-control cell co-cultures (Fig. 3D). This demonstrates that DUX4 is able to activate target genes in murine nuclei initially devoid of any DUX4 protein. This co-culture system allows us to conclude that after transcription of DUX4 in a very small number

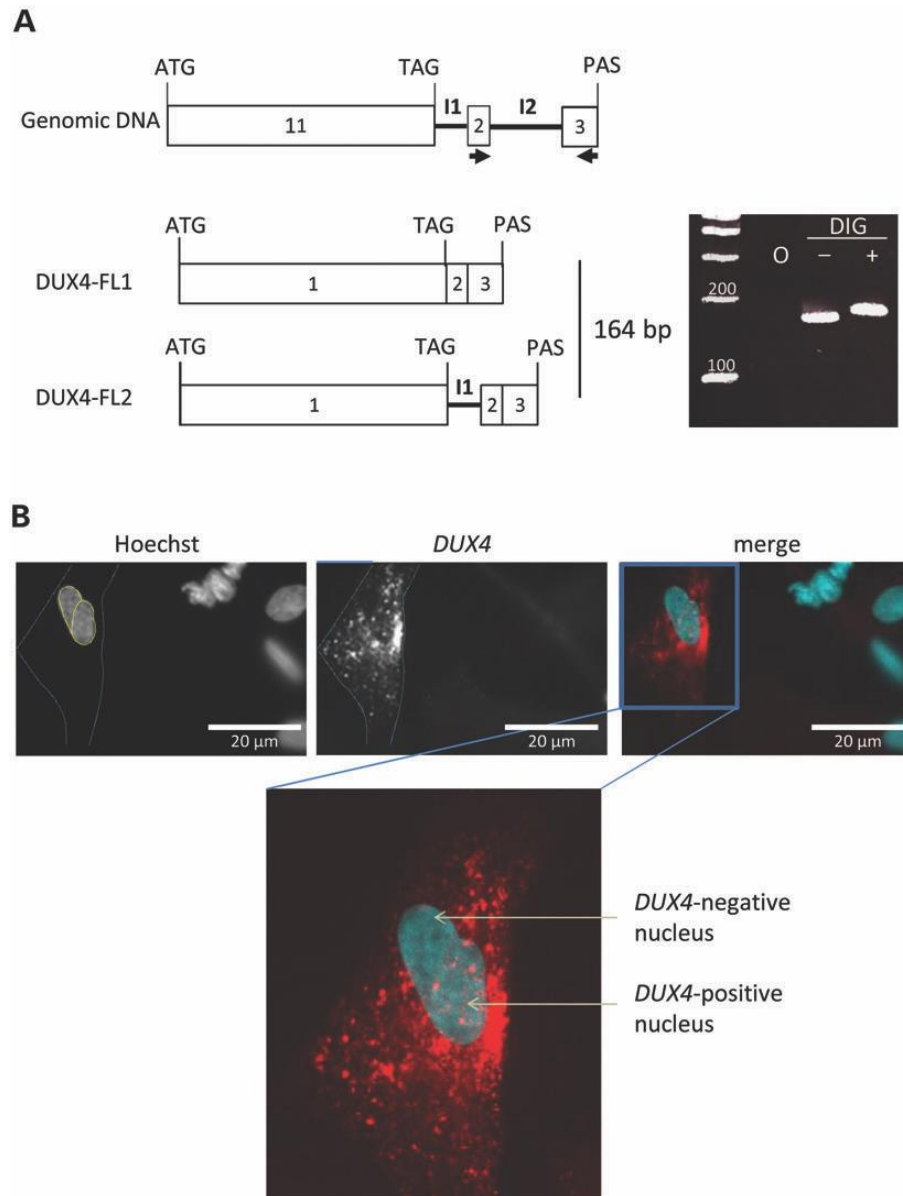


Figure 1. DUX4 mRNA is expressed in few nuclei within myotube. RNA-FISH against DUX4-FL mRNA was performed on FSHD myotubes at Day 5 of differentiation. (A) The probe was synthesized using the PCR DIG Probe Synthesis Kit (right panel) and recognizes all DUX4-FL isoforms (left panel). The primers used as probes are represented by black arrows and generates a PCR amplicon of 164 bp without the DIG. (B) DUX4 RNA-FISH in FSHD myotubes. In the merge picture, DUX4 labeling is in red, nuclei (Hoescht 33342, 1/1000, Dako fluorescent mounting) in blue. The nucleus and cell borders are, respectively, in yellow and blue and the Zoom: myotube containing 1 DUX4-positive and 1 DUX4-negative nuclei.

of human nuclei within myotubes, DUX4 mRNA is translated in the cytoplasm and the resulting proteins spread into both human and murine nuclei where they activate genes downstream of DUX4. Thus, the molecular signature of FSHD is spread throughout myonuclei within myotubes, which can explain how a very rare toxic mRNA, present in only 1 in 1000 nucleus (9) can participate efficiently to the FSHD pathophysiology.

DM1 nuclei induced splicing defects in murine nuclei in heterokaryons

Consequences of nuclear protein spreading were also analyzed in DM1. We investigated by using co-cultures whether splicing

factors produced by murine nuclei, can be titrated by DM1 CUGexp-RNAs trapped in neighboring human nuclei within the same myotube and consequently, induce splicing misregulation events in the murine nuclei. Co-cultures between DM1, control or FSHD human cell lines and C2C12 cells were performed and harvested at Day 5 of differentiation. After reverse transcription (RT), splicing of Serca1 exon 22 and Bin1 exon 11, that were shown to be mis-spliced in DM1 muscle cells (15,19), were analyzed. Using human (SERCA1) or murine (Serca1) specific primers, we observed that in heterokaryons formed between C2C12 and DM1 human cells, splicing of both SERCA1 and Serca1 exon 22 are impaired in the presence of DM1 cells when compared with co-cultures with control or

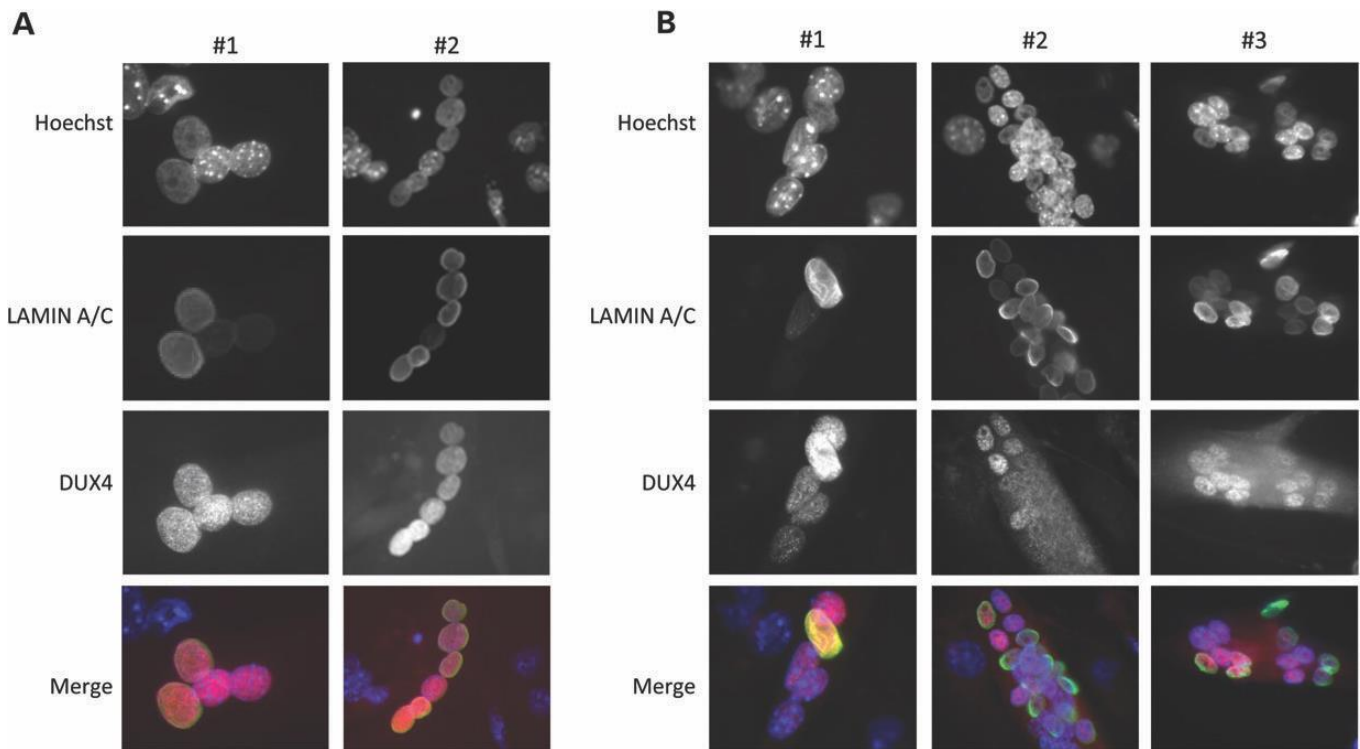


Figure 2. DUX4 is present in murine nuclei in heterokaryons. (A) Co-culture experiments were performed between human FSHD cells and murine C2C12 cells. At Day 5 of differentiation, cells were fixed and immunostainings were performed: DUX4 was labeled with E5.5 antibody (red), human nuclei were stained with human-specific LAMIN A/C (green) and nuclei were counterstained using Hoechst 33342 (blue). (B) Examples of DUX4 gradient in heterokaryons. # represents independent fields.

FSHD cells (Fig. 4A). Indeed, the percentage of inclusion of human SERCA1 exon 22 is 53% in heterokaryons formed between C2C12 and control cells and 20% when the heterokaryons are formed between C2C12 cells and DM1 cells. Similarly, in the co-cultures performed between C2C12 and non-DM1 cells, the percentage of inclusion of mouse Serca1 is 49% but only 30% in the C2C12-DM1 co-cultures (Fig. 4B). Splicing of Serca1 exon 22 is thus impaired in both human and murine nuclei by the presence of the mutated DMPK mRNA. Splicing of BIN1/Bin1 is also altered in both human and murine nuclei in the co-cultures with DM1 cells. In the absence of the mutated CUGexp-RNAs, the percentage of inclusion of murine Bin1 exon 11 represents 72%, whereas in the presence of the mutated mRNA, this percentage drops to 47%. Finally, we observed that Mbln1 exon 7 splicing is also impaired in the DM1-C2C12 co-cultures when compared with FSHD or control co-cultures (Fig. 4C).

These results show that within a syncytium, pre-mRNAs transcribed in murine nuclei surrounded by DM1 nuclei develop splicing defects, suggesting a trans-dominant effect of human CUGexp-RNA on murine splicing and the spreading of the DM1 phenotype to other nuclei within the myotubes.

CUGexp-RNA aggregates are localized only in human nuclei

Finally, we investigated whether the splicing defects we observed were related to the presence of the CUGexp-RNAs within the murine nuclei. A RNA-FISH directed against CUG expansion

was performed and revealed that CUGexp-RNA aggregates are only localized in human nuclei, confirming that the mis-splicing events observed in the murine nuclei are due to the trans-dominant depletion of murine splicing factors by the CUGexp-RNAs present in human nuclei and not to the presence of the human CUGexp-RNAs in aggregates within the murine nuclei (Fig. 4D). As for FSHD, the molecular signature of DM1 is spread from human to murine nuclei throughout myotubes.

DISCUSSION

The fact that muscle is a multinucleated organ has often been seen as a protection against the effect of some mutations. For instance and despite the random X chromosome inactivation, females heterozygous for mutations in the dystrophin gene will not develop Duchenne Muscular Dystrophy thanks to the synthesis of DYSTROPHIN by the healthy allele in 50% nuclei and the diffusion of the functional protein throughout the myotube at the sarcolemma. Another example is the accumulation of mutations in the mitochondrial genome of fibers can be somewhat limited by the fusion of satellite cells, which do not accumulate these mutations at the same rate (20,21). This phenomenon is not true for all muscle proteins and depends on the size of the nuclear domain that might impair intracellular movement of proteins (22). The spreading of nuclear proteins has been described several years ago using heterokaryons (3,23). However, the consequences on pathophysiology and therapies for muscular diseases had not been investigated.

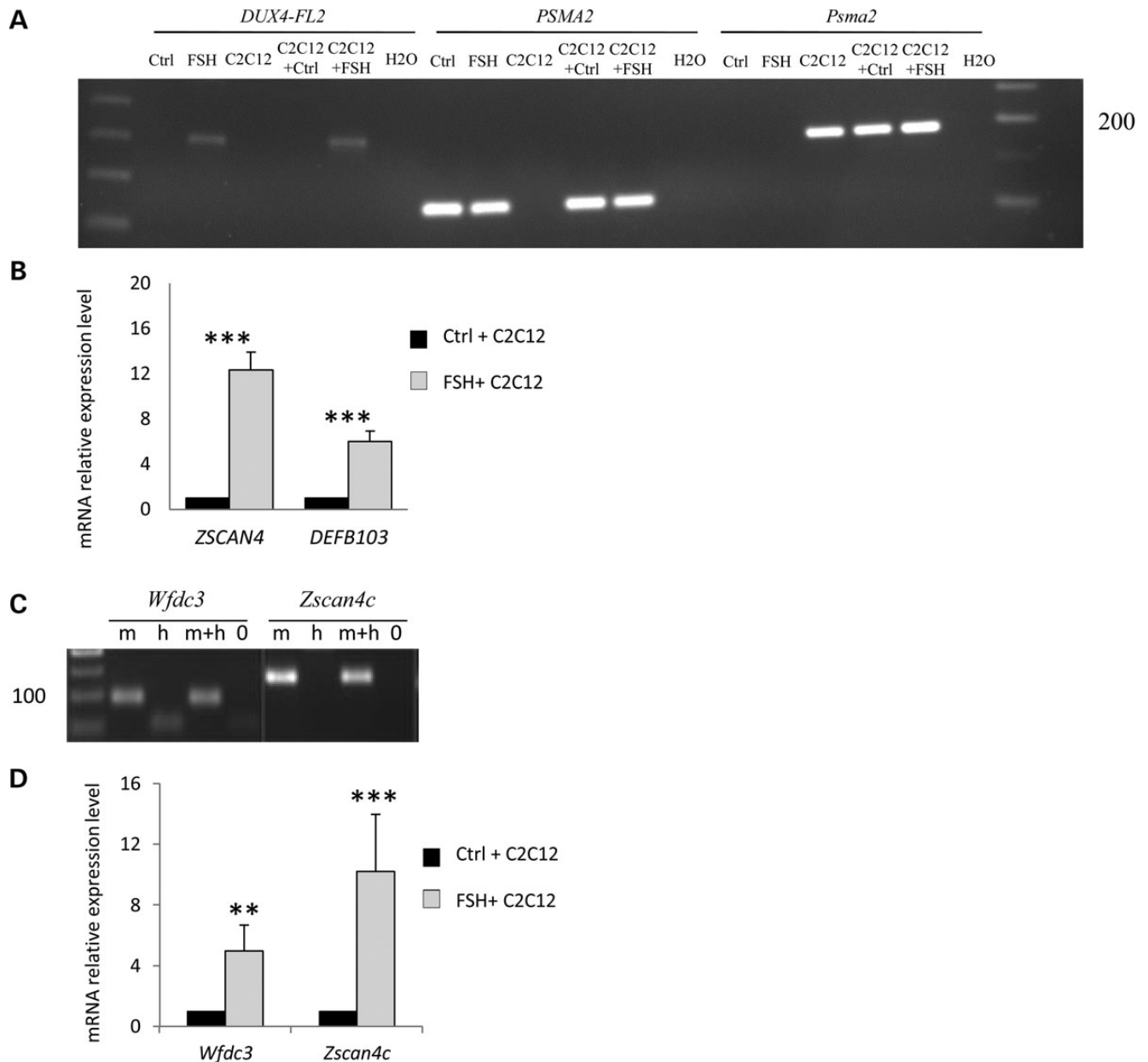


Figure 3. DUX4 activates human and murine genes in heterokaryons. (A) A RT-PCR performed on either DUX4 or PSMA2 (housekeeping) on the different co-cultures or on C2C12 and FSHD cells alone at Day 5 of differentiation shows that DUX4-FL2 is only expressed in the presence on FSHD cells. (B) ZSCAN4 and DEFB103, 2 human genes downstream of DUX4, are activated only when C2C12 and human FSHD cell lines are co-cultured. (C) Species specificity of the primers used. RT-PCR was performed on either murine (m), human (h), or a 1 : 1 mix of human-murine (m+h) cDNAs. With the different primers, the presence of an amplified product is observed only when the murine cDNA is present, which demonstrates the murine specificity of the primers. (D) Expression levels of murine genes downstream of DUX4 are measured by qPCR in the co-cultures. PSMA2 and Psm2 are used as normalizers for human and murine mRNAs, respectively. Data represent mean \pm standard deviation from at least three independent experiments. ***P \leq 0.001 by Student's t-test.

In this study, we investigated whether nuclear protein spreading could be evidenced in myotubes. Concerning FSHD, one of the key questions about the pathophysiology of this disease is how such a rare event the transcription of DUX4 in only a few nuclei can trigger the worsening of entire muscle fibers? Even though the hypothesis was raised that within myotubes, DUX4 may be transcribed in few nuclei but the resulting protein may spread to other nuclei inducing a cascade of abnormal gene activations, it was never demonstrated (11,24). Here, we show the spreading of the DUX4 protein from one FSHD nucleus to

neighboring nuclei, thus demonstrating that spreading of nuclear proteins can exist and in additional are functional since genes downstream of DUX4 are activated. This spreading may explain that, when investigating the presence of the protein DUX4, a decrease in signal intensity is observed in consecutive nuclei within myotubes. This consecutive gradient of DUX4 intensity is observed in both human and murine nuclei. Indeed, as all proteins, DUX4 is translated in the cytoplasm and is redirected to the nuclei by three monopartite nuclear localization signals (25). The probability to reach one specific nucleus may depend on the distance

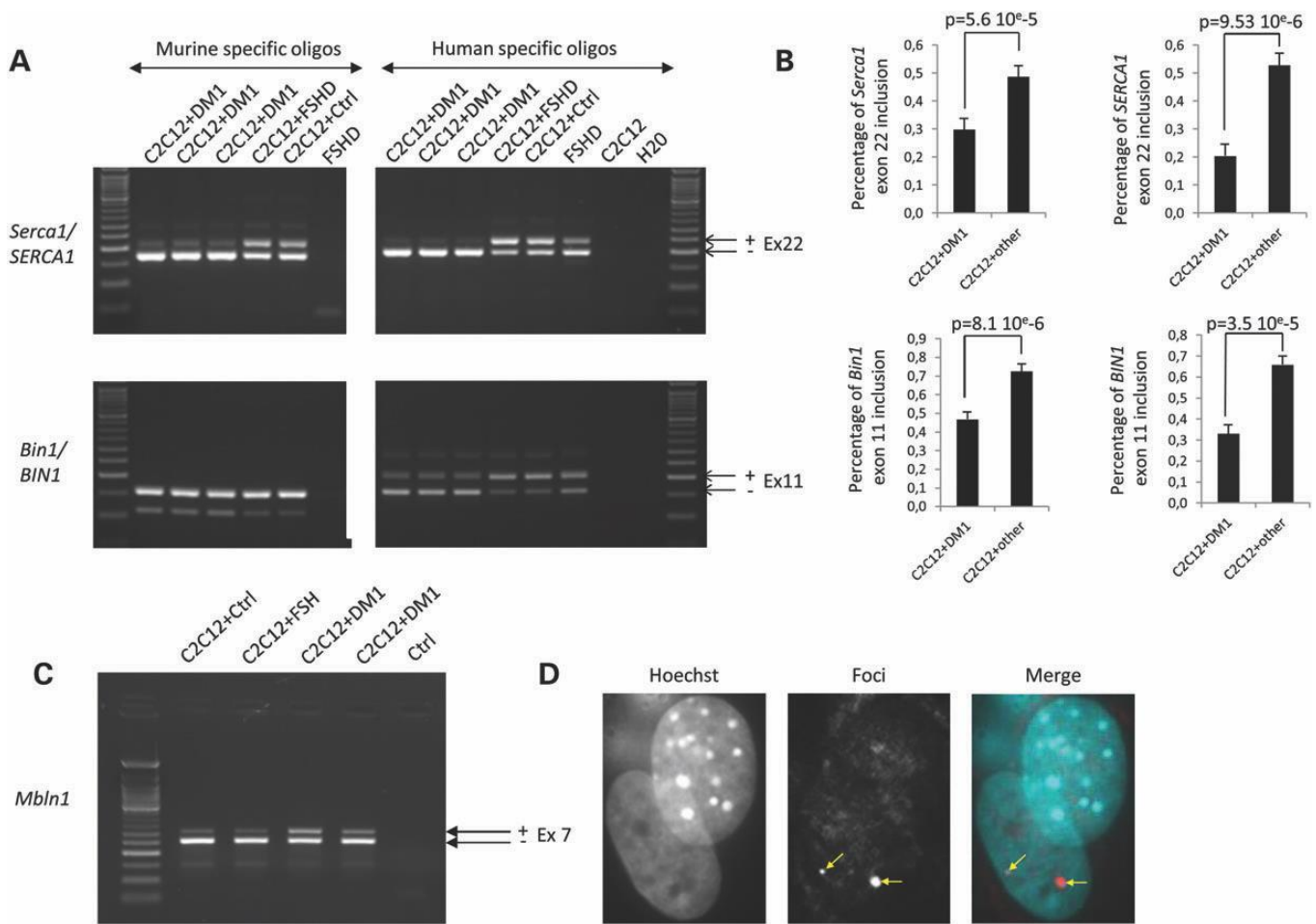


Figure 4. DM1 phenotype spread in heterokaryons. (A) Representative RT–PCR analysis of either endogenous human BIN1 or murine Bin1 (BIN1/Bin1) and SERCA1/ Serca1 from myotubes formed in co-cultures between murine C2C12 cells and either DM1, FSHD or control cells. (B) The percentage of inclusion of either murine (left panel) or human (right panel) Serca1 exon22 or Bin1 exon11 as indicated was measured on 11 samples and Student's t-test was performed. All data represent mean \pm standard error of the mean. (C) Murine splicing of Mbln1 exon 7 is impaired in the presence of DM1 cells. Representative RT–PCR analysis of endogenous murine Mbln1 from myotubes formed in co-cultures between murine C2C12 cells and either DM1, FSHD or control cells. (D) RNA-FISH analysis showing the localization of CUGexp-RNA aggregates (red) in the nuclei. Punctuated mouse and uniformly stained human nuclei are clearly distinguished by Hoechst 33342 (blue).

and on the proteins involved in the nuclear import. The longest distance we measured between one human DUX4-positive nucleus and one murine DUX4-positive nucleus was 51 μ m. However, the nucleus of a myoblast moves rapidly after fusion towards the central myotube nuclei. It is thus difficult to conclude that DUX4 can migrate as far as 51 μ m because two nuclei can be adjacent at a given time and then rapidly migrate at the two extremities of the myotubes (26). However, it is possible that in vivo, during fiber repair, the nucleus newly incorporated in muscle fiber is briefly close enough to a DUX4 expressing nucleus to be 'contaminated' and thus spread DUX4-associated molecular phenotype.

Another interesting point raised about DUX4 propagation within myotube is whether or not the transactivation of either human or murine gene is due to a strong activation from a very limited number of nuclei or a weak activation from many nuclei (24). It has been very recently published that the number of DUX4-positive nuclei is 6–8-fold increased when myoblasts are differentiated using 20% knock out serum replacer (KSOR) instead of classical conditions (serum depletion+ITS) (10). In parallel, the activation of the genes downstream of DUX4 is

increased in the KSOR cultures, showing that the amplification of the expression of downstream targets of DUX4 is correlated to the number of DUX4-positive nuclei. Consequently, the effect of the spreading is more likely linked to a weak activation from most of the DUX4-positive nuclei rather than a strong activation from one adjacent nucleus.

Concerning DM1, we show in DM1-C2C12 heterokaryons that nuclear CUGexp-RNAs transcribed by DM1 nuclei sequester splicing factors produced by C2C12 nuclei, inducing murine pre-mRNA mis-splicing. In both FSHD and DM1 heterokaryons, a spreading of the pathological phenotype from one nucleus to others is thus confirmed. Such diffusion may be detrimental only for gain-of-function diseases, and similar mechanisms may occur in other muscular diseases involving nuclear protein dysfunctions such as oculopharyngeal muscular dystrophy (OPMD) in which mutated PABPN1 is aggregated with numerous nuclear proteins (27). The physiological consequences of this spreading may depend on the size of the D4Z4 repeats for FSHD, the level of expansion of the CTG trinucleotide repeat for DM1 or on the expansion of the trinucleotide (GCN) repeat

in the first exon of PABPN1 for OPMD. Moreover, several other parameters such as the half-life of the mRNA or proteins involved, their expression levels and their functions may also be important.

Another aspect of the consequences of such diffusion is a potential impact on therapeutic approaches. Questions such as how many nuclei will have to be repaired by gene transfer or how many normal cells have to be grafted to counteract the pathological spreading needs to be investigated, and the answers may be patient specific. However, several studies have now started targeting the nuclear machinery to avoid viral replication in the nucleus or improve drug delivery (28) and similar approaches could be developed for FSHD to avoid DUX4 dissemination.

Finally, multinucleated cells are present in several human organs among them the liver (hepatocytes), the brain (Purkinje cells) and bones (osteoclasts) where similar spreading may occur and may impact therapy efficiency. For instance, somatic cell fusion events between bone marrow derived-cells and Purkinje cells have been proposed to counteract spinocerebellar ataxia type 1, but since the pathological mechanism of this disease involves expansion of the polyglutamine tract in ATXN1 which induces a loss of RORa from Purkinje cell nuclei (29), such strategies may not be successful.

In conclusion, we show that spreading of nuclear proteins or mRNA participates to the dissemination of a pathological molecular phenotype within multinucleated fibers, and may represent a novel mechanism to explain pathophysiology. This should be taken into account in the design of new therapeutic strategies.

MATERIALS AND METHODS

Cell culture

Immortalized FSHD myoblast clones were generated from one mosaic patient as previously described (30) and the immortalized DM1 myoblasts used in this study carries 2500 CTG repeats in the 3' UTR of the DMPK gene. Human clones were cultivated in proliferation medium [4 vols of DMEM, 1 vol of 199 medium, FBS 20%, gentamycin 50 mg/ml (Life technologies, Saint Aubin, France)] supplemented with insulin 5 mg/ml dexametasone 0.2 mg/ml, b-FGF 0.5 ng/ml, hEGF 5 ng/ml and fetuine 25 mg/ml. Murine C2C12 were cultivated in DMEM, FBS 10%, gentamycin 50 mg/ml. Differentiation medium for both human and murine cells was composed of DMEM supplemented with insulin (10 mg/ml). For FSHD experiments, heterokaryon formation was obtained by mixing clones and C2C12 at ratio 4 : 1 (1.5×10^4 and 3.75×10^4 cells per cm^2 , respectively) on gelatin-coated plates or slides in proliferation medium. Two days later, proliferation medium was replaced by differentiation medium. The fusion efficacy was improved by treating cells with PEG 1000 at a concentration of 50% (W/V) in DMEM pH 7.4 for 60 s at 37°C according to the method as previously described (23). Cells were carefully rinsed and cytosine arabinoside (Ara-C, 10^{-5} M, Sigma-Aldrich) was added to reduce the proportion of unfused cells. Five days after the induction of differentiation, the cells were trypsinized and filtered using pre-separation filters of 30 μm (Miltenyi Biotech, Paris, France) to selectively retain myotubes. For DM1 heterokaryon experiments, a ratio 4 : 1 between DM1/C2C12 cells has been used. The myoblasts

Table 1. Sequences of the primers

Targeted gene	Primer	Sequence
DUX4-FL2	DUX4_FL2_fw	CCTGGGATTCTGCCTTCTA
	DUX4_FL2_rev	AGCCAGAAATTCACGGAAGA
B2M	B2M_fw	CTCTCTTTCTGGCCTGGAGG
	B2M_rev	TGCTGGATGACGTGAGTAAACC
PSMA2	PSMA2_fw	CATTCAGCCCGTCTGGTAAA
	PSMA2_rev	GTTTTTCTCAGTTGCTAATACCACA
Psm2	Psm2_fw	AGAGCGCGGTTACAGCTTC
	Psm2_rev	CTCCACCTTGTAACACTCCTT
DEFB103	DEFB103_fw	GAGGATCCATTATCTTCTGTTTGC
	DEFB103_rev	AGGCAGCTGAGCACAGCA
Wfdc3	Wfdc3_fw	GGTAGCTGCAGGAGAGCACG
	Wfdc3_rev	CTGGGTCAGGATTCTCTC
ZSCAN4	ZSCAN4_fw	GTGGCCACTGCAATGACAA
	ZSCAN4_rev	AGCTTCCTGTCCCTGCATGT
Zscan4c	Zscan4c_fw	GATTATTGGCCACAGGACAAG
	Zscan4c_rev	TCAGGGTGCTGTTCTTCTG
SERCA1	SERCA1_fw	ATCTTCAAGCTCCGGGCCCT
	SERCA1_rev	CAGCTCTGCCTGAAGATGTG
Serca1	Serca1_fw	GCTCATGGTCTCAAGATCTCAC
	Serca1_rev	GGGTCACTGCCTCAGCTTTC
BIN1	BIN1_fw	AGAACCTCAATGATGTGCTGG
	BIN1_rev	TCGTGGTTGACTCTGATCTCGG
Bin1	Bin1_fw	GGGAGCAACACCTTCACAGT
	Bin1_rev	GATCTCAGGGGTAGCAGCAG
Mb1n1	Mb1n1_fw	CCATCCTCCCGCACACC
	Mb1n1_rev	TGGTGGGAGAAATGCTGTATGC

were plated on gelatin-coated plates and differentiation was induced 2 days later. Five days after triggering differentiation, cells were harvested and filtered to retain myotubes.

RNA extraction, PCR and real-time PCR

Trizol was directly added on filters and RNA extraction was performed according to the manufacturer protocol (Life technologies, Saint Aubin, France). RNA concentration was determined using a nanodrop ND-1000 spectrophotometer (Thermo Scientific, Wilmington, USA). The RT was carried out on 1 mg of total RNA with Roche Transcriptor First Strand cDNA Synthesis Kit (Roche, Meylan, France) at 60°C for 50 min with 1 ml of oligo dT and 2 ml of random hexamers in a final volume of 20 μl .

Quantitative PCRs were performed in a final volume of 9 μl with 0.4 μl of RT product, 0.18 μl of each forward and reverse primers 20 pmol/ml (Table 1), and 4.5 μl of SYBR[®]Green mastermix 2 \times (Roche, Meylan, France). The qPCR was run in triplicates on a LightCycler 480 Real-Time PCR System (Roche, Meylan, France). The qPCR cycling conditions were 94°C for 5 min, followed by 50 cycles at 95°C for 30 s and 60°C for 15 s and 72°C for 15 s. The PCR for BIN1/Bin1 and SERCA1/Serca1 were performed as previously described for BIN1 (31). PSMA2 and Psm2 were used, respectively, as human and murine normalizer as determined by the geNorm software.

Immunofluorescence staining

For co-immunostaining of DUX4 and LAMIN A/C, cells were fixed in 2% paraformaldehyde for 10 min at room temperature and washed twice with PBS 1 \times . Cells were permeabilized with

1% Triton X-100 (Sigma) in PBS for 10 min at room temperature with gentle rocking. After blocking non-specific sites with PBS-FBS 2%, E5.5 DUX4 antibody directed against the C-terminal region of DUX4 (dilution 1:10) and antibody against human LAMIN A/C (clone JOL2, mouse IgG1; 1:100; Abcam, Cambridge, MA, USA) were diluted in PBS-FBS 2% and incubated with the cells overnight at 4°C. After three washes with PBS-0.5% Triton X-100, the cells were incubated with the secondary antibodies [goat anti rabbit IgGH+L–Alexa 488 (1/200) and goat anti mouse IgG1–Alexa 555 (1/200)] for 1 h in the dark. The cells were washed three times with PBS-0.5% Triton X-100 and incubated 5 min with Hoescht 33342 (1/1000, Thermo scientific, Courtaboeuf, France) before being mounted on microscope slides using Dako fluorescent mounting medium (Dako, Glostrup, Denmark).

RNA-FISH

The DUX4 probe was synthesized using the PCR DIG Probe Synthesis Kit (Roche, Meylan, France) with a 1 : 6 ratio of DIG-dUTP. The primers are DUX4_FL2_fw and DUX4_FL2_rev allowing a DUX4-specific amplification of 164 bp. The PCR was performed in a final volume of 50 µl using the taq platinum and 1 µl of cDNA, according to a manufacturer protocol. The probe was purified using the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel, Hoerdt, France). The RNA-FISH was performed on immortalized FSHD myotubes at Day 5 of differentiation. The cells were fixed for 10 min in PFA 4%, rinsed with PBS and incubated with hybridization buffer (formamide 40%, 2×SSC, 0.2% BSA)+20 ng of DUX4 probe for 90 min at 50°C. The cells were then washed and incubated with PBS-tween 0.1% for 30 min at 55°C, and non-specific binding was blocked with PBS-FBS 2% for 30 min at room temperature. Cells were then incubated with anti-digoxigenin-Rhodamine Fab Fragments (dilution 1/200, Roche diagnostics) for 1 h in the dark. After three washes with PBS, the cells were incubated 5 min with Hoescht 33342 and mounted on microscope slides using Dako fluorescent mounting medium.

The probe used to detect DM1 foci was a 2' OMe oligo containing the sequence (CAG)₇-Cy3. The cells were fixed in PFA 4% for 10 min, washed in PBS and incubated in hybridization buffer+probe (dilution 1/1000) for 90 min at 37°C. The cells were then washed and incubated in PBS-tween 0.1% for 30 min at 42°C, then rinsed in PBS before being incubated with Hoescht 33342 and mounted on microscope slides using Dako fluorescent mounting.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at HMG online.

Conflict of Interest statement. None declared.

FUNDING

This work was supported by the Association Française contre les Myopathies (AFM-Téléthon, France), UPMC-Emergence 2010 (to J.D.), FSHD Global research foundation Ltd (to M.F.), the FSH society (FSHS-22012-03 to V.M.), and the Agence

Nationale de la Recherche (FSHDecipher, ANR-13-BSV1-0004 to J.D.). The authors acknowledge the platform of the Institut de Myologie and in particular Kamel Mamchaoui for immortalization of human cells. The E5.5 DUX4 antibody was a kind gift from Dr S. Tapscott and pci-neo and pci-DUX4 a gift from Dr A. Belayew.

REFERENCES

- Denetclaw, W.F. Jr, Bi, G., Pham, D.V. and Steinhardt, R.A. (1993) Heterokaryon myotubes with normal mouse and Duchenne nuclei exhibit sarcolemmal dystrophin staining and efficient intracellular free calcium control. *Mol. Biol. Cell.*, 4, 963–972.
- Blau, H.M., Chiu, C.P., Pavlath, G.K. and Webster, C. (1985) Muscle gene expression in heterokaryons. *Adv Exp Med Biol*, 182, 231–247.
- Ralston, E. and Hall, Z.W. (1989) Transfer of a protein encoded by a single nucleus to nearby nuclei in multinucleated myotubes. *Science*, 244, 1066–1069.
- van Deutekom, J.C., Bakker, E., Lemmers, R.J., van der Wielen, M.J., Bik, E., Hofker, M.H., Padberg, G.W. and Frants, R.R. (1996) Evidence for subtelomeric exchange of 3.3 kb tandemly repeated units between chromosomes 4q35 and 10q26: implications for genetic counselling and etiology of FSHD1. *Hum. Mol. Genet.*, 5, 1997–2003.
- Lemmers, R.J., Tawil, R., Petek, L.M., Balog, J., Block, G.J., Santen, G.W., Amell, A.M., van der Vliet, P.J., Almomani, R., Straasheijm, K.R. et al. (2012) Digenic inheritance of an SMCHD1 mutation and an FSHD-permissive D4Z4 allele causes facioscapulohumeral muscular dystrophy type 2. *Nat. Genet.*, 44, 1370–1374.
- Sacconi, S., Lemmers, R.J., Balog, J., van der Vliet, P.J., Lahaut, P., van Nieuwenhuizen, M.P., Straasheijm, K.R., Debipersad, R.D., Vos-Versteeg, M., Salvati, L. et al. (2013) The FSHD2 gene SMCHD1 is a modifier of disease severity in families affected by FSHD1. *Am. J. Hum. Genet.*, 93, 744–751.
- van der Maarel, S.M., Miller, D.G., Tawil, R., Filippova, G.N. and Tapscott, S.J. (2012) Facioscapulohumeral muscular dystrophy: consequences of chromatin relaxation. *Curr. Opin. Neurol.*, 25, 614–620.
- Gabriels, J., Beckers, M.C., Ding, H., De Vriese, A., Plaisance, S., van der Maarel, S.M., Padberg, G.W., Frants, R.R., Hewitt, J.E., Collen, D. et al. (1999) Nucleotide sequence of the partially deleted D4Z4 locus in a patient with FSHD identifies a putative gene within each 3.3 kb element. *Gene*, 236, 25–32.
- Snider, L., Geng, L.N., Lemmers, R.J., Kyba, M., Ware, C.B., Nelson, A.M., Tawil, R., Filippova, G.N., van der Maarel, S.M., Tapscott, S.J. et al. (2010) Facioscapulohumeral dystrophy: incomplete suppression of a retrotransposed gene. *PLoS Genet.*, 6, e1001181.
- Block, G.J., Narayanan, D., Amell, A.M., Petek, L.M., Davidson, K.C., Bird, T.D., Tawil, R., Moon, R.T. and Miller, D.G. (2013) Wnt/beta-catenin signaling suppresses DUX4 expression and prevents apoptosis of FSHD muscle cells. *Hum. Mole. Genet.*, 22, 390–396.
- Tassin, A., Laoudj-Chenivresse, D., Vanderplanck, C., Barro, M., Charron, S., Anseau, E., Chen, Y.W., Mercier, J., Coppee, F. and Belayew, A. (2012) DUX4 expression in FSHD muscle cells: how could such a rare protein cause a myopathy? *J. Cell. Mol. Med.*, 17, 76–89.
- Ferreboeuf, M., Mariot, V., Bessieres, B., Vasiljevic, A., Attie-Bitach, T., Collardeau, S., Morere, J., Roche, S., Magdinier, F., Robin-Ducellier, J. et al. (2014) DUX4 and DUX4 downstream target genes are expressed in fetal FSHD muscles. *Hum. Mole. Genet.*, 23, 171–181.
- Klein, A.F., Gasnier, E. and Furling, D. (2011) Gain of RNA function in pathological cases: focus on myotonic dystrophy. *Biochimie*, 93, 2006–2012.
- Francois, V., Klein, A.F., Beley, C., Jollet, A., Lemerrier, C., Garcia, L. and Furling, D. (2011) Selective silencing of mutated mRNAs in DM1 by using modified hU7-siRNAs. *Nat. Struct. Mol. Biol.*, 18, 85–87.
- Fugier, C., Klein, A.F., Hammer, C., Vassilopoulos, S., Ivarsson, Y., Toussaint, A., Tosch, V., Vignaud, A., Ferry, A., Messaddeq, N. et al. (2011) Misregulated alternative splicing of BIN1 is associated with T tubule alterations and muscle weakness in myotonic dystrophy. *Nat. Med.*, 17, 720–725.
- Geng, L.N., Tyler, A.E. and Tapscott, S.J. (2011) Immunodetection of human double homeobox 4. *Hybridoma (Larchmt)*, 30, 125–130.

17. Geng, L.N., Yao, Z., Snider, L., Fong, A.P., Cech, J.N., Young, J.M., van der Maarel, S.M., Ruzzo, W.L., Gentleman, R.C., Tawil, R. et al. (2012) DUX4 activates germline genes, retroelements, and immune mediators: implications for facioscapulohumeral dystrophy. *Dev. Cell*, 22, 38–51.
18. Krom, Y.D., Thijssen, P.E., Young, J.M., den Hamer, B., Balog, J., Yao, Z., Maves, L., Snider, L., Knopp, P., Zammit, P.S. et al. (2013) Intrinsic epigenetic regulation of the D4Z4 macrosatellite repeat in a transgenic mouse model for FSHD. *PLoS Genet.*, 9, e1003415.
19. Kimura, T., Nakamori, M., Lueck, J.D., Pouliquin, P., Aoike, F., Fujimura, H., Dirksen, R.T., Takahashi, M.P., Dulhunty, A.F. and Sakoda, S. (2005) Altered mRNA splicing of the skeletal muscle ryanodine receptor and sarcoplasmic/endoplasmic reticulum Ca^{2+} -ATPase in myotonic dystrophy type 1. *Hum. Mole. Genet.*, 14, 2189–2200.
20. Clark, K.M., Bindoff, L.A., Lightowlers, R.N., Andrews, R.M., Griffiths, P.G., Johnson, M.A., Brierley, E.J. and Turnbull, D.M. (1997) Reversal of a mitochondrial DNA defect in human skeletal muscle. *Nat. Genet.*, 16, 222–224.
21. Spendiff, S., Reza, M., Murphy, J.L., Gorman, G., Blakely, E.L., Taylor, R.W., Horvath, R., Campbell, G., Newman, J., Lochmuller, H. et al. (2013) Mitochondrial DNA deletions in muscle satellite cells: implications for therapies. *Hum. Mole. Genet.*, 22, 4739–4747.
22. Pavlath, G.K., Rich, K., Webster, S.G. and Blau, H.M. (1989) Localization of muscle gene products in nuclear domains. *Nature*, 337, 570–573.
23. Blau, H.M., Chiu, C.P. and Webster, C. (1983) Cytoplasmic activation of human nuclear genes in stable heterocaryons. *Cell*, 32, 1171–1180.
24. Ehrlich, M. and Lacey, M. (2012) Deciphering transcription dysregulation in FSH muscular dystrophy. *J. Hum. Genet.*, 57, 477–484.
25. Corona, E.D., Jacquelin, D., Gatica, L. and Rosa, A.L. (2013) Multiple protein domains contribute to nuclear import and cell toxicity of DUX4, a candidate pathogenic protein for facioscapulohumeral muscular dystrophy. *PLoS One*, 8, e75614.
26. Cadot, B., Gache, V., Vasyutina, E., Falcone, S., Birchmeier, C. and Gomes, E.R. (2012) Nuclear movement during myotube formation is microtubule and dynein dependent and is regulated by Cdc42, Par6 and Par3. *EMBO Rep.*, 13, 741–749.
27. Banerjee, A., Apponi, L.H., Pavlath, G.K. and Corbett, A.H. (2013) PABPN1: molecular function and muscle disease. *FEBS J.*, 280, 4230–4250.
28. Chahine, M.N. and Pierce, G.N. (2009) Therapeutic targeting of nuclear protein import in pathological cell conditions. *Pharmacol. Rev.*, 61, 358–372.
29. Serra, H.G., Duvick, L., Zu, T., Carlson, K., Stevens, S., Jorgensen, N., Lysholm, A., Burright, E., Zoghbi, H.Y., Clark, H.B. et al. (2006) ROR α -mediated Purkinje cell development determines disease severity in adult SCA1 mice. *Cell*, 127, 697–708.
30. Krom, Y.D., Dumonceaux, J., Mamchaoui, K., den Hamer, B., Mariot, V., Negroni, E., Geng, L.N., Martin, N., Tawil, R., Tapscott, S.J. et al. (2012) Generation of isogenic D4Z4 contracted and noncontracted immortal muscle cell clones from a mosaic patient: a cellular model for FSHD. *Am. J. Pathol.*, 181, 1387–1401.
31. Brouqsault, N., Morere, J., Gaillard, M.C., Dumonceaux, J., Torrents, J., Salort-Campana, E., Maues de Paula, A., Bartoli, M., Fernandez, C., Chesnais, A.L. et al. (2013) Dysregulation of 4q35- and muscle-specific genes in fetuses with a short D4Z4 array linked to facio-scapulo-humeral dystrophy. *Hum. Mole. Genet.*, 22, 4206–4214.

AUTRES RESULTATS :
Un micro ARN issu de l'intron 1 de
DUX4 ?

1. Introduction

1.1 Les micros ARN

Les récents progrès réalisés en matière de compréhension de la régulation de l'expression ont permis l'identification et la caractérisation de nouveaux acteurs : les micros ARN (miARN). Les miARN sont de petits ARN capables de se fixer par complémentarité à un ou plusieurs transcrits cibles et d'induire un blocage post-transcriptionnel de leur expression. C'est en 1993 que le premier miARN a été identifié et associé à une fonction biologique. En effet, chez *C. elegans*, un petit ARN non codant de 22 nucléotides, appelé lin-4, est capable de réguler la traduction de l'ARN lin-14 en se fixant à sa 3'UTR (Lee et al., 1993; Wightman et al., 1993) et jouant un rôle important dans la régulation du développement de *C. elegans*. Depuis cette découverte, un nombre croissant de miRNA a été identifié et cela chez la plupart des espèces (Lagos-Quintana et al., 2001; Lau et al., 2001; Lee and Ambros, 2001).

1.2 Fonction des micros ARN

Les miARN sont des ARN non codants simple brin dont la taille est comprise entre 20 et 25 nucléotides. A l'heure actuelle, environ un millier de micro ARN ont été décrits. De par la très forte conservation de leurs séquences au cours de l'évolution, environ 60% des gènes seraient soumis à une régulation par les miARN. D'une manière générale, les miARN ciblent les régions 3'UTR des transcrits (Bartel, 2004, 2009; Bentwich et al., 2005; Friedman et al., 2009; Lim et al., 2003) et provoquent soit la dégradation de l'ARNm, soit la répression traductionnelle (Bartel, 2004; Lim et al., 2005).

2. Biogenèse des micros ARN

2.1 Transcription des miARN

Les miARN peuvent être issus soit de régions intergéniques, souvent regroupés en cluster (Lagos-Quintana et al., 2001; Lau et al., 2001; Lee et al., 2002; Mourelatos et al., 2002), et donc sous la dépendance de leur propre promoteur ou alors être d'origine intragénique et donc dépendant de l'expression d'un transcrit donné (Lagos-Quintana et al., 2001; Lau et al., 2001; Lee and Ambros, 2001; Lee et al., 2004). Chez l'Homme, 50 à 80% des miARN sont localisés dans des régions introniques de gènes codants, la majorité dans la portion centrale de l'intron (Rodriguez et al., 2004; Kim and Kim, 2007; Morlando et al., 2008).

Les miARN sont transcrits sous forme d'un long ARN appelé pri-miARN pouvant contenir plusieurs pré-miARN (intergénique), ou être inclus dans un transcrit codant (intragénique) (**Figure 27**). Ce transcrit primaire possède une cape en 5' ainsi qu'une queue poly(A) en 3', démontrant que celui-ci est produit à partir d'une ARN polymérase de type 2 (Tam, 2001; Aukerman and Sakai, 2003).

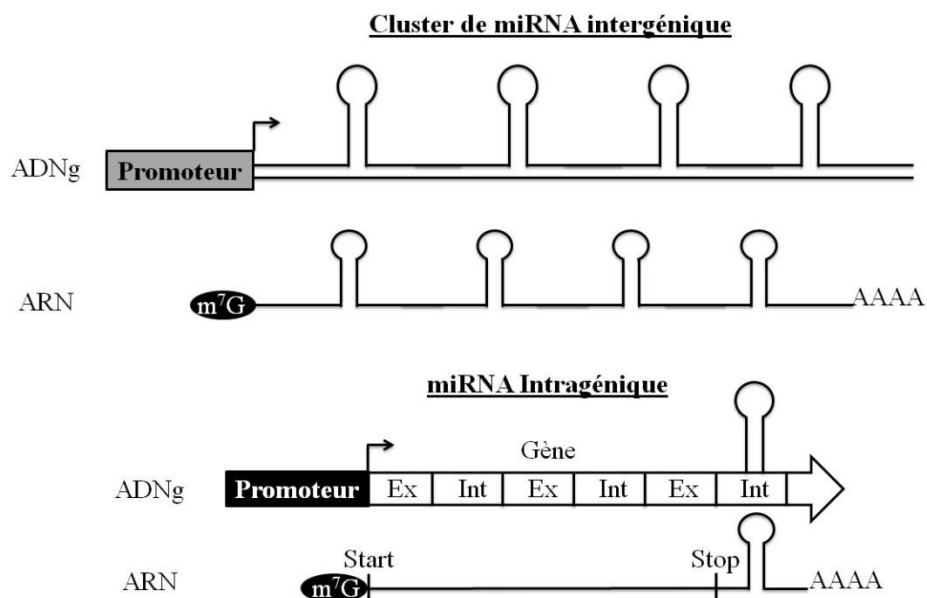


Figure 27 : Localisation et transcription des miARN intergéniques et intragéniques :

Les miARN intergénique sont regroupés sous forme de cluster et leur transcription est dépendante de leur propre promoteur

Les miARN intragéniques sont eux majoritairement localisés dans les 3'UTR de gènes, leur transcription est donc dépendante du gène hôte.

2.2 Clivage par la machinerie DROSHA-DGCR8 et export

Un seul pri-miARN peut contenir jusqu'à 6 miARN précurseurs, possédant une structure secondaire dite en « épingle à cheveux » appelés pre-miARN. Le pri-miARN est reconnu et clivé par le complexe « microprocesseur » constitué d'une protéine de liaison à l'ARN double brin (DGCR8) et d'une enzyme RNase de type 3 (DROSHA) (Denli et al., 2004; Gregory et al., 2004; Han et al., 2004; Landthaler et al., 2004; Lee et al., 2003). DGCR8 est le premier acteur dans le clivage du pri-miARN car il guide DROSHA sur la structure en « épingle à cheveux ». DGCR8 se fixe sur la portion double brin située à la jonction entre le simple brin du pri-miARN et la structure en épingle à cheveux de celui-ci. DROSHA peut alors s'associer à DGCR8 et venir cliver l'ARN à environ 11 paires de bases de la jonction (Han et al., 2006) (**Figure 28**). Le produit ainsi généré fait environ 60 à 70 nucléotides et est appelé pré-miARN. Il conserve sa structure en tige boucle et est constitué d'une extrémité 5'phosphate et d'une extrémité 3' hydroxyl avec 2 nucléotides sortants permettant sa maturation ultérieure.

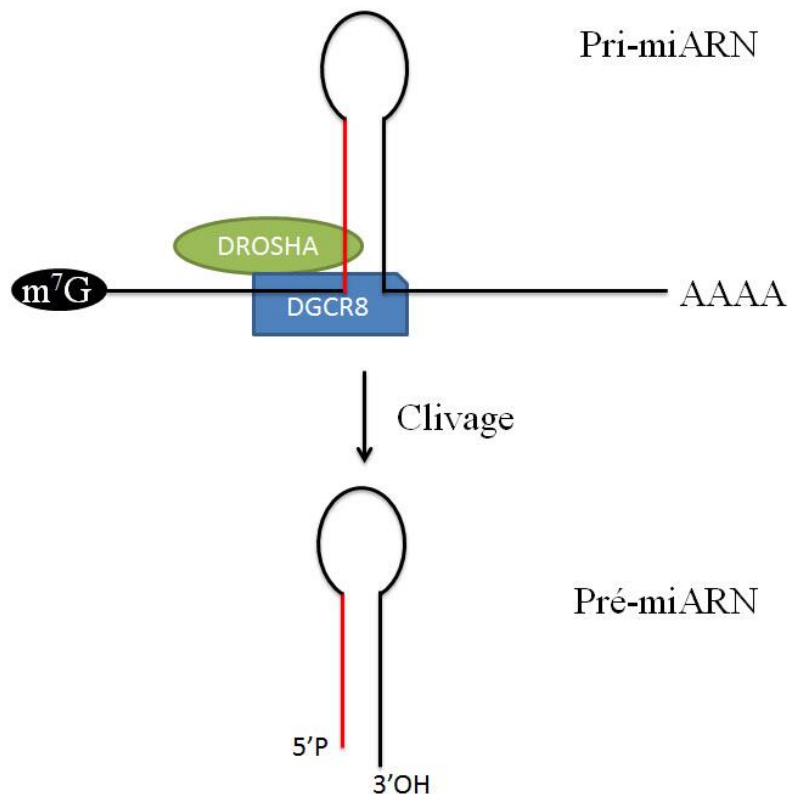


Figure 28: Maturation des pri-miARN en pré-miARN par le complexe microprocesseur : DGCR8 fixe la structure ARN double brin permettant ainsi la fixation de DROSHA qui clive

la structure tige boucle de l'ARN hôte libérant ainsi le pré-miARN possédant une extrémité 5'P rentrante nécessaire à son export vers le cytoplasme.

Le pré-miARN est ensuite exporté du noyau vers le cytoplasme à l'aide d'une protéine cargo appartenant à la famille des karyophérines : l'exportin 5 associée à la protéine Ran-GTP (Ras-related nuclear protein fixant le GTP) indispensable à l'export nucléaire. Il semble que cette protéine soit majoritairement utilisée pour le transport des pré-miARN (Yi et al., 2003; Gwizdek et al., 2004). La reconnaissance du pré-miARN par l'exportin 5 n'est pas séquence dépendante, en effet, il a été montré que l'association passe par une reconnaissance de la structure ARN double brin, composant la tige du pré-miARN, de manière Ran-GTP dépendante (Bohnsack et al., 2004; Zeng and Cullen, 2004). Enfin l'extrémité 3' sortant du pré-miARN est indispensable à un export efficace car elle favorise la reconnaissance du pré-miARN par l'exportin 5.

2.3 Maturation cytoplasmique médiée par DICER

Une fois exporté dans le cytoplasme, le pré-miARN subit une nouvelle étape de maturation médiée par une RNase cytoplasmique de type 3 appelée DICER, générant un duplexe micro ARN de 19 à 23 nucléotides. DICER, comme les autres RNase III, clive en laissant deux nucléotides libres à l'extrémité 3' de son substrat servant à la sélection de l'un des 2 brins. En effet, en fonction de la stabilité thermodynamique duplex (Khvorova et al., 2003; Krol et al., 2004; Schwarz et al., 2003), l'un des 2 brins est sélectionné (brin guide) et chargé dans le complexe micro ARN-RISC (miRISC : micro ARN induced silencing complexe) tandis que le second est dégradé (brin star). Néanmoins, il arrive que le brin star soit assez stable pour être chargé dans le miRISC et jouer le rôle de miRNA (Okamura et al., 2008).

Le complexe RISC est composé d'une molécule de miARN, d'une protéine de type Argonaute (AGO) et d'une protéine de type GW182. Les protéines Argonautes possèdent quatre domaines distincts : les domaines N-terminal, Paz, Mid et PIWI. Chez les mammifères, il existe 4 protéines AGO, AGO1 à AGO4, mais seul le domaine PIWI de AGO2 possède une activité enzymatique de type RNase-H like. Le complexe miRISC identifie l'ARNm par complémentarité avec le miARN et conduit soit au clivage du transcrit (complémentarité

miARN/mARN cible parfaite) soit à l'inhibition de la traduction de l'ARNm cible (complémentarité imparfaite) (Bartel, 2004; Lim et al., 2005).

Ainsi, chez l'Homme, les micros ARN sont capables de cibler un ARNm de façon séquence-spécifique. Une complémentarité parfaite n'est pas nécessaire à l'activité du miARN. En effet, seule la région allant du 2^{ème} aux 8^{ème} nucléotides (cotés 5') nécessite un appariement parfait. Cette région, appelée séquence « seed », est donc critique pour la reconnaissance de la cible et une mutation de l'un de ces nucléotides diminue significativement l'efficacité d'action du miRNA. La région 3' dont l'appariement n'est pas indispensable permet néanmoins de stabiliser le duplex miARN-ARNcible. La taille de la région « seed » étant réduite (7 nucléotides) plusieurs microARN peuvent cibler des centaines d'ARNm différents et un même ARN peut être la cible de plusieurs miARN permettant une régulation négative efficace par effets additifs (Bartel, 2004).

3. Les micros ARN introniques

Contrairement à ce qui était connu sur les miARN, il apparaîtrait que la majorité des miARN humains soient localisés dans des séquences introniques de régions codantes ou non (Rodriguez et al., 2004). Ainsi, l'expression de ces miARN est directement liée à la transcription du gène hôte. Les miARN introniques sont dépendants de DROSHA/DGCR8 qui clivent la structure tige boucle sur l'ARN pré messager, mais ce mécanisme se déroule en synergie avec le spliceosome (machinerie d'épissage)(Kim and Kim, 2007) (**Figure 29**).

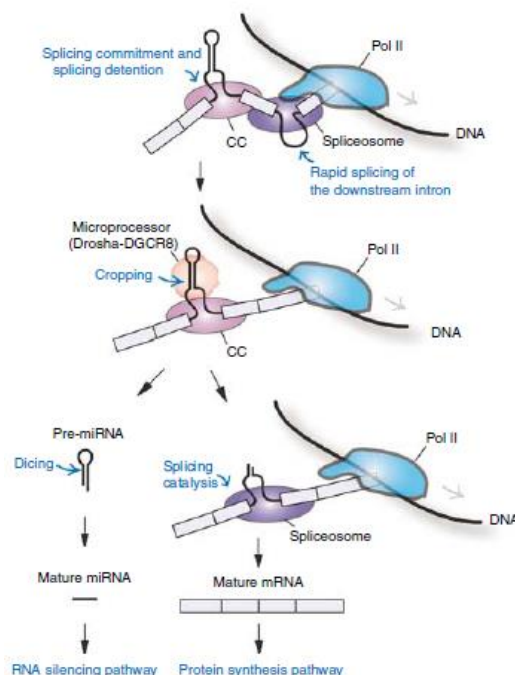


Figure 29 : Un modèle du processing des micro ARN intronique (Kim and Kim, 2007).

4. Cas particulier des mirtrons

Contrairement aux miARN canoniques où la biogenèse est liée au clivage du pre-miARN par DROSHA/DGCR8, une seconde voie de production indépendante vis-à-vis du complexe DROSHA /DGCR8 a été identifiée. Dans cette voie, les structures en tige boucle (similaires aux précurseurs des miARN) adoptées sur un intron sont libérées lors de la maturation du pré ARN messenger. Le lasso, produit d'épissage contenant l'intron, peut alors être débranché par une enzyme appelée Ldbr (Lariat debranching enzyme) générant ainsi une structure de type pré-miARN appelé mirtron. Classiquement, les extrémités 5' ou 3' des mirtrons correspondent aux sites donneurs ou accepteurs d'épissage (Ruby et al., 2007), néanmoins, il est possible que les extrémités 3' ou 5' ne soient pas les bornes du mirtron mature. Dans ces cas de figures, après débranchage, le pré-mirtron subit une étape de maturation où les extrémités sont dégradées (Flynt et al., 2010). Les mirtrons rejoignent alors la même voie de biogenèse des miARN (export, maturation par DICER, prise en charge par le microprocesseur) (**Figure 30**).

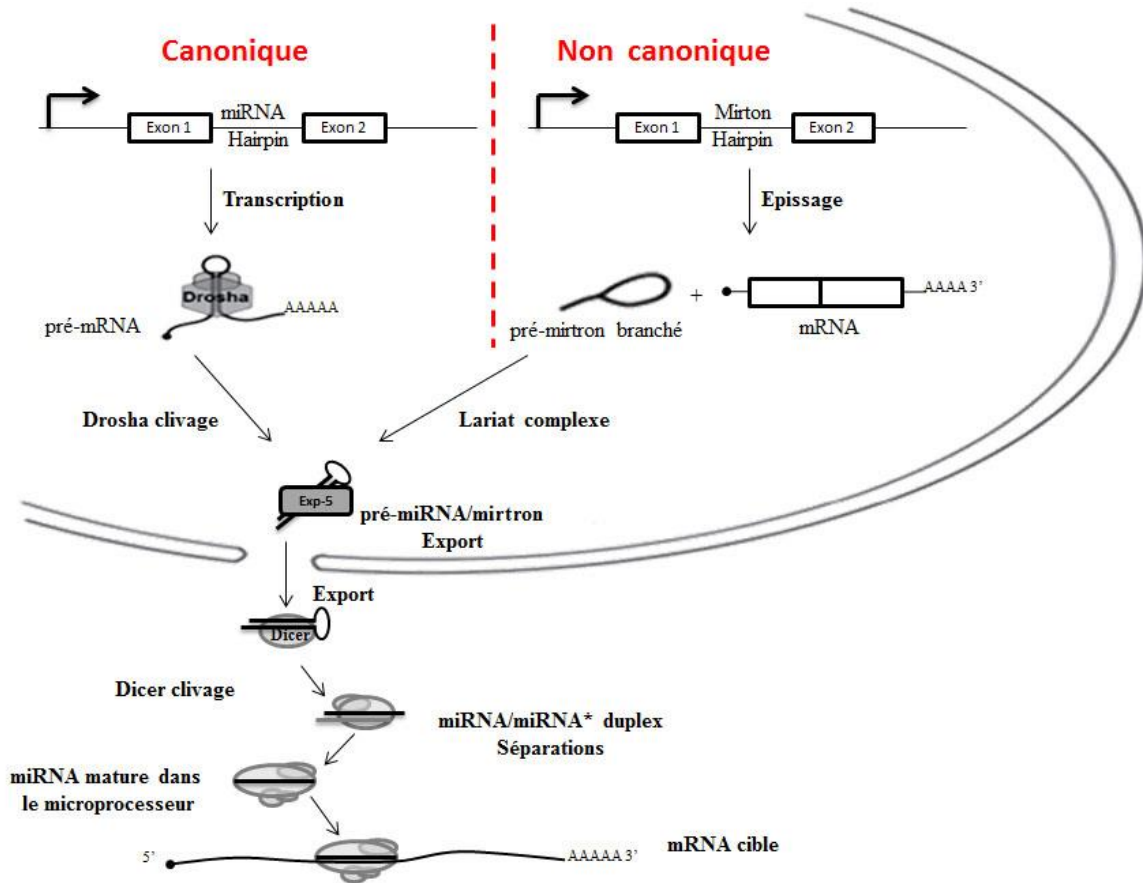


Figure 30 : Schéma récapitulatif des voies de biogenèse des miARN canonique et non canonique.

5. Résultats

5.1 Identification d'un petit ARN non codant par bioinformatique

L'ARNm DUX4 contient 2 introns situés dans sa région 3'UTR donc non codante. Or, le premier intron est soumis à un épissage alternatif. Cet épissage n'influençant pas la séquence protéique, le rôle de cet intron est posé. Puisqu'un grand nombre d'ARN non codants sont issus d'introns, nous avons dans un premier temps réalisé une analyse bioinformatique prédictive du premier intron de DUX4 afin de déterminer si sa séquence nucléotidique pouvait donner naissance à un ARN non codant. L'utilisation du site bioinformatique <http://bioinfo3.noble.org/> a permis de prédire que les bases 12 à 32 du premier intron pouvaient correspondre à un miARN. De manière intéressante, l'analyse bioinformatique a prédit un grand nombre de cibles putatives dont certaines sont directement

liées au muscle comme le dystroglycane 1 (dystrophin-associated glycoprotein 1, impliqué dans l'architecture du muscle) ou encore un canal calcium (impliqué dans la contraction musculaire). La FSHD étant une pathologie musculaire, l'identification de cibles putatives jouant un rôle dans le muscle font de ce miARN un candidat très intéressant dans la physiopathologie de la maladie.

L'une des caractéristiques d'un miARN est la structure en « épingle à cheveux » de son précurseur. La modélisation de la structure de l'intron 1 via mFold a prédit une conformation en « tige boucle » en accord avec l'architecture classique d'un précurseur de miARN (**Figure 31**).

Au vue de ces résultats, il apparait donc plausible qu'un miARN soit produit à partir de l'intron 1 de DUX4.

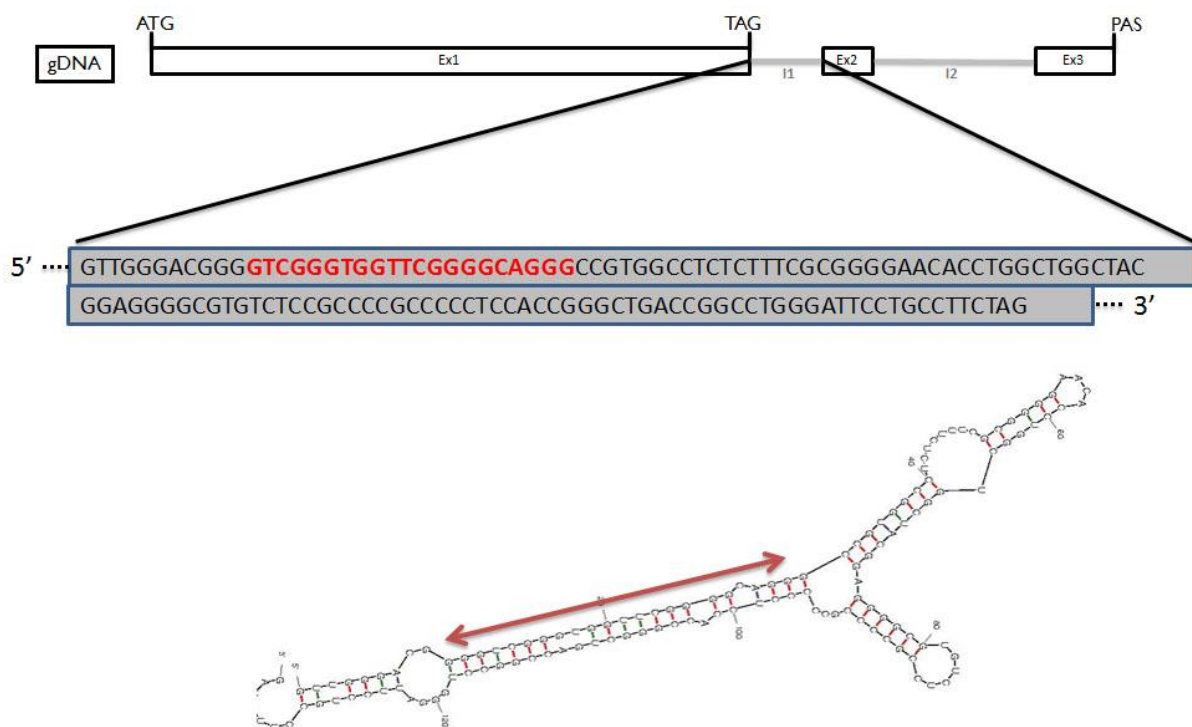


Figure 31 : Position du miARN prédit sur DUX4 et structuration de l'intron :

L'intron 1 de DUX4 (136 pb) situé dans la 3'UTR pourrait générer un miARN (en rouge) et adopterait une structure compatible avec celle observée chez les précurseurs des miARN.

5.2 Mise en évidence d'un ARN non codant par Northern Blot

Afin de déterminer si l'intron 1 de DUX4 est capable de générer un miARN, nous avons réalisé des analyses par Northern Blot sur une extraction de petits ARN issus de quadriceps de patients contrôle et FSHD (fœtus de 16 semaines de développement). Trois sondes différentes ont été utilisées, la première, d'une longueur de 21 nucléotides, est l'exacte complémentarité du miARN mature prédit, la seconde sonde est quant à elle strictement complémentaire du miARN mature prédit ainsi que d'une partie de la tige boucle (brin guide). Enfin la troisième sonde est strictement complémentaire du brin antisens (brin star) et d'une partie de la tige boucle. Les deux premières sondes sont donc capables de s'hybrider avec les différents produits de maturation du miARN tandis que la troisième sonde ne donnera aucun signal pour le miARN mature à condition que le brin star ne soit pas assez stable pour donner également naissance à un miARN (**Figure 32**).

3 sondes utilisées :

- Sonde 1 : Strictement complémentaire du miRNA mature
- Sonde 2 : Complémentaire du brin formant le miRNA mature (brin guide) et d'une partie de la tige boucle
- Sonde 3 : Complémentaire du brin antisens (brin star) et d'une partie de la tige boucle

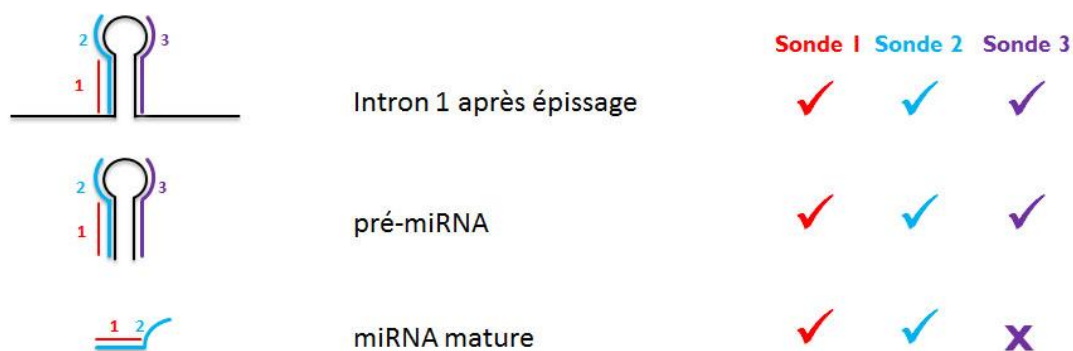


Figure 32 : Récapitulatif des sondes utilisées pour détecter le miARN intronique par Northern Blot.

Les différentes isoformes de DUX4 sont détectées dès le stade fœtal, de plus, l'expression est la plus forte vers le 5^{ème} jour de différenciation, c'est pourquoi, nous avons réalisé un premier essai sur des myoblastes fœtaux à J5 de différenciation.

Grâce au Northern Blot, nous avons pu confirmer le résultat issu de l'analyse bio-informatique. Les 3 produits issus des étapes de formation du miRNA sont visibles avec les sondes 1 et 2 permettant donc de valider sa présence. De plus, l'absence de marquage pour le miRNA mature avec la sonde 3 (violet) permet d'affirmer que seul le brin guide est stabilisé et permet la production d'un miRNA mature.

De manière intéressante, le pri-miARN détecté correspond en taille à celle de l'intron 1 en entier, suggérant que le produit d'épissage de l'intron 1 est débranché et peut donc être à l'origine du pre-miARN. Bien que l'ARNm DUX4 soit faiblement exprimé (1 noyau sur 1000) le signal correspondant à l'intron 1 de DUX4 en entier est très important. De plus, les différents fragments sont détectés à la fois dans le quadriceps control et FSHD et cela à une intensité similaire (**Figure 33**).

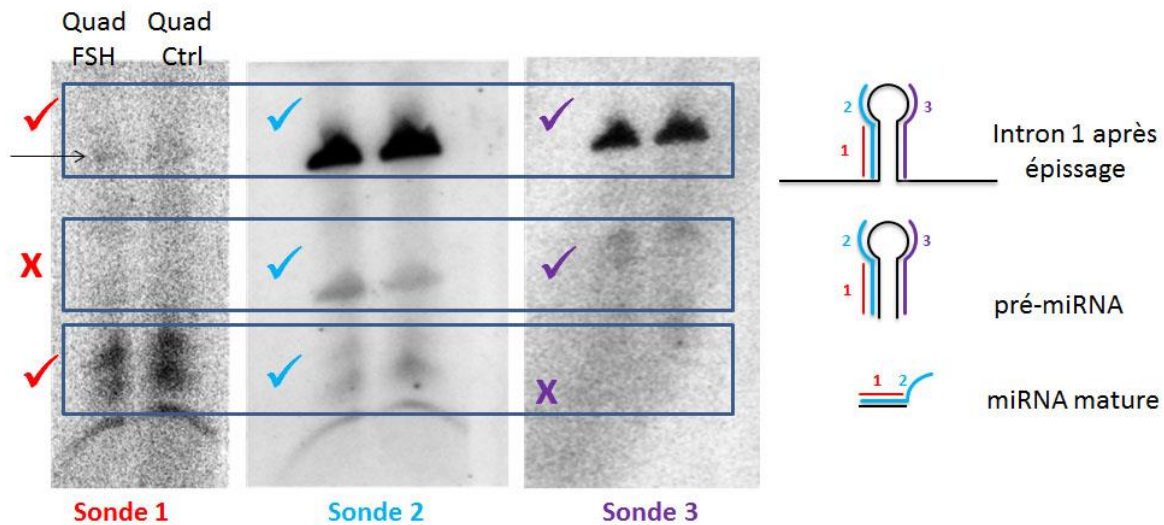


Figure 33 : Northern Blot sur les cellules fœtales contrôle et FSHD.

Ces résultats suggèrent que l'intron 1 de DUX4 est épissé et peut être débranché (voie de biogenèse on canonique) afin de générer un miARN mature. De plus, sa détection à la fois dans les cellules FSHD et Contrôle suggère que le produit d'épissage ne provient pas exclusivement du chromosome 4. En effet, les répétitions D4Z4 des chromosomes 4 et 10 partagent 98% homologie de séquence, il est alors possible que les fragments détectés soient produits par le chromosome 10. Cette conclusion serait en accord avec une étude de 2009 montrant que de petits fragments d'ARN étaient produits à partir des répétitions D4Z4, mais

compte tenu de la forte homologie entre les répétitions du chromosome 4 et 10 il n'est pas possible de distinguer leurs origines (Snider et al., 2009).

6. Discussion et perspectives

6.1 Résultats à confirmer

Bien que ces premiers résultats soient encourageants, ils restent à être confirmés. En effet, nous n'avons pour le moment réalisé qu'un seul test et nous ne pouvons donc pas confirmer que les fragments détectés soient spécifiques. Néanmoins, l'utilisation d'un jeu de trois sondes différentes fournit chacun des résultats attendus suggérant que les fragments sur lesquels s'hybrident les sondes ne sont pas aspécifiques.

6.2 Caractérisation de la voie de production du micro ARN

Si la présence du miARN dans l'intron 1 de DUX4 est confirmée, il sera nécessaire de mieux le caractériser notamment en déterminant quelle voie de production est utilisée. Pour déterminer si ce miARN est DROSHA/DGCR8 dépendant ou non, nous utiliserons une stratégie d'interférence à l'ARN visant à bloquer DROSHA et nous vérifierons l'impact sur l'expression du miARN par northern blot. Ainsi, si nous détectons encore la production du miARN mature, nous pourrions conclure que la production est DROSHA/DGCR8 indépendante et que le miARN appartient donc à la classe des mirtrons. Nous utiliserons également cette stratégie pour l'enzyme DICER qui bloquera la production du miARN mature qui nous permettra d'affirmer que le fragment détecté par Northern Blot est pris en charge par la machinerie des miARN et aboutit à la production d'un miARN mature.

Si la production du miARN est DROSHA/DGCR8 indépendante (classe des mirtrons), nous utiliserons un shARN ciblant DBR1, bloquant ainsi l'étape clé permettant le débranchage du lasso après épissage, ce qui nous permettra de confirmer l'appartenance à la classe des mirtrons (**Figure 34**).

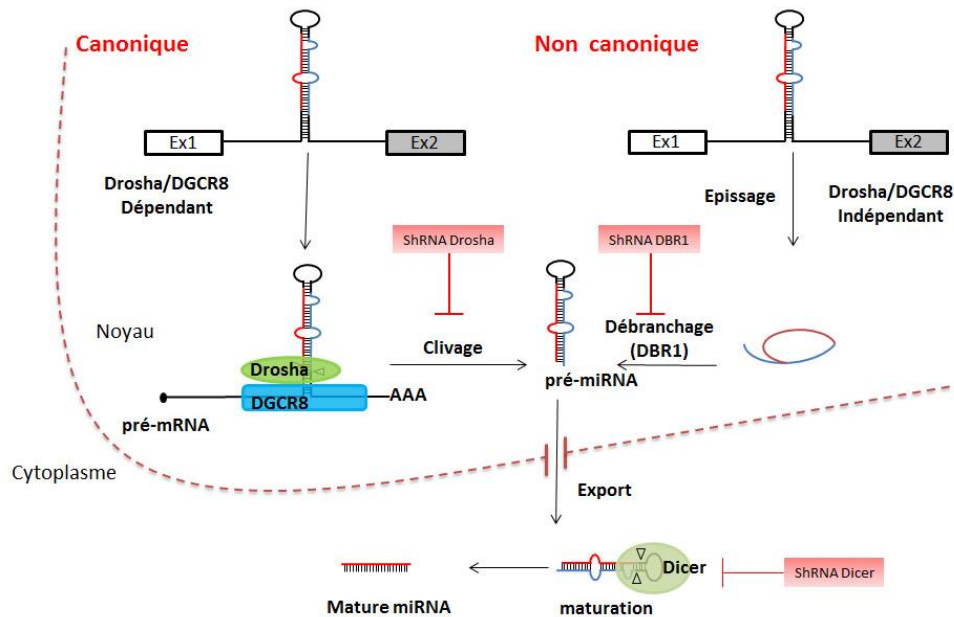


Figure 34: Stratégie visant à déterminer la classe du miARN détecté par utilisation de shARN ciblant les voies clés de maturation.

L'utilisation d'un shARN contre DROSHA permettra de bloquer uniquement la voie canonique tandis que l'utilisation d'un shARN contre DICER bloquera la maturation des pré-miARN en miARN matures des deux voies.

6.3 Vérification de la fonctionnalité du miARN

Pour déterminer si l'intron 1 de DUX4 est bien à l'origine de la production du miARN, il sera cloné en phase dans l'ORF de la GFP (cassant ainsi la phase) tout en conservant les sites donneurs et accepteurs d'épissage. Les cellules GFP positives seront donc celles ayant correctement épissées l'intron. Pour savoir si cet épissage produit bien un miARN nous réaliserons un Northern Blot sur ces cellules. Enfin, la production du miARN ne suffit pas à conclure de sa fonctionnalité, ainsi nous pourrons réaliser une éponge à miARN en ciblant artificiellement la mCherry par insertion dans sa 3'UTR de la séquence cible du miARN, ainsi une diminution de la fluorescence de la mCherry sera synonyme d'extinction par le miARN produit par notre construction GFP-Intron1 DUX4 (**Figure 35**).

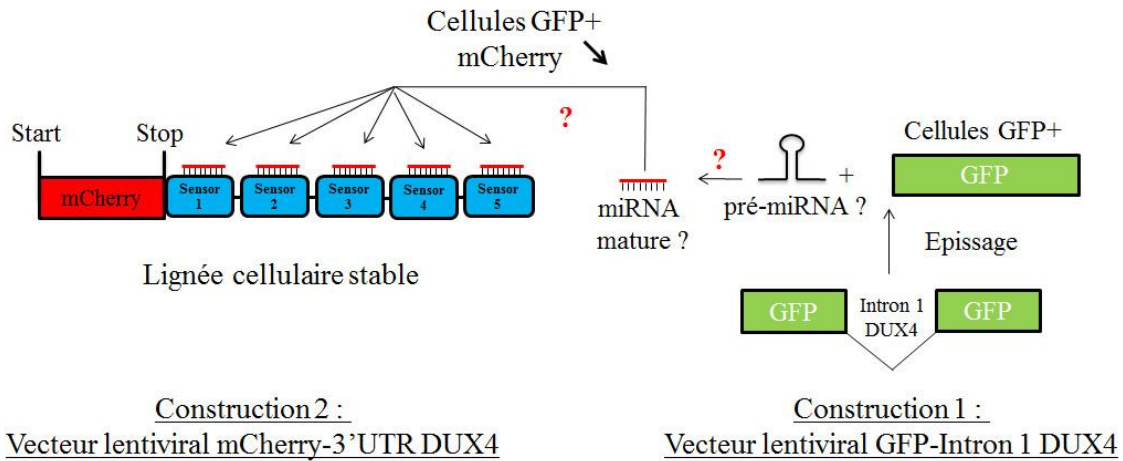


Figure 35: Stratégie visant à déterminer si le miARN est fonctionnel :

L'épissage de l'intron 1 permettra une restauration de la phase de la GFP et donc la fluorescence. Si un miARN est bien produit à partir de l'intron, il pourra alors réguler l'expression de la mCherry via les séquences sensor insérées dans la 3'UTR de la mCherry.

6.4 Le miARN peut-il venir réguler DUX4 ?

L'une des hypothèses serait que l'épissage de l'intron 1 de DUX4, permettant la production du miARN mature, auto-régule DUX4. Puisque la majorité des miRNA cible les régions 3'UTR, selon la même stratégie que celle décrite précédemment, cette région de DUX4 sera elle aussi clonée dans la 3'UTR de la mCherry, afin de déterminer si le miRNA produit peut agir sur son propre transcrit. Ainsi, si le miARN peut fixer la 3'UTR de DUX4, alors il induira une diminution du niveau d'expression de la mCherry (**Figure 36**).

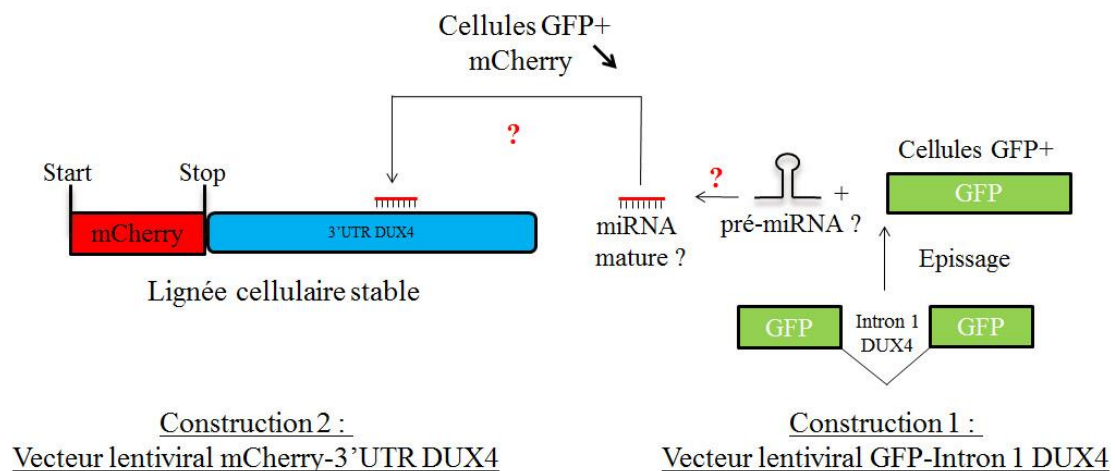


Figure 36: Stratégie visant à déterminer si le miARN peut interagir avec la 3'UTR de DUX4.

Le rôle de l'épissage alternatif dans la 3'UTR de DUX4 est inconnu. Si nous confirmons qu'un miARN est bien produit à partir de l'intron 1 de DUX4, alors ce résultat fournira une nouvelle voie d'étude sur la fonction de DUX4. Toutefois, nos résultats préliminaires semblent montrer que ce miARN est présent à la fois dans les cellules FSHD et Contrôles suggérant donc que ce miARN n'est potentiellement pas impliqué dans la pathologie.

7. Matériels et méthodes

7.1 Extraction et séparation des ARN

L'ARN total des cellules est extrait en utilisant du Trizol conformément au protocole du fournisseur (Life technologies, Saint Aubin, France). La quantité d'ARN extrait est mesurée par spectrophotométrie à l'aide d'un nanodrop 1000 (Thermo Scientific, Wilmington, USA). L'appareil réalise une mesure à 260 nm correspondant au pic d'absorbance de l'ARN. L'intensité est directement proportionnelle à la quantité d'ARN et il est alors possible de déterminer la quantité totale d'ARN.

Les petits et les longs ARN sont ensuite purifiés sur colonne à l'aide du kit mirVana™ miRNA Isolation Kit conformément au protocole du fournisseur (Life technologies, Saint Aubin, France). Classiquement, sur 50 µg d'ARN total déposé sur la colonne, approximativement 10% de petites ARN sont récupérés.

7.2 Préparation et migration sur gel de polyacrylamide

Un gel 10% Acrylamide/bis acrylamide (19:1) en condition dénaturante (urée) est préparé en MOPS-NaOH 20mM (pH 7). Avant de déposer les échantillons, le gel est préchauffé durant 2 heures sous un courant de 450 V (18 mA) en tampon de migration (MOPS-NaOH 20mM). 5 µg de petits ARN sont alors mélangés avec du bleu de charge (ARN) contenant du formamide, incubés 5 min à 95°C afin de dénaturer les structures secondaires, puis déposés sur le gel.

7.3 Transfert sur membrane de nitrocellulose et cross link chimique

Le gel est dans un premier temps incubé dans un bain de BET afin de vérifier la qualité de migrations des échantillons. La membrane de nitrocellulose membrane (GeneScreen Plus® Hybridization Transfer Membrane, PerkinElmer) est ensuite imprégnée de tampon de transfert (0,5X of Tris-Borate-EDTA) puis placée directement en contact du gel. Trois à quatre couches de papier watman imprégnées de tampon de transfert sont ensuite placées de part et d'autre du gel et de la membrane et le tout est inséré dans une cassette de transfert permettant de maintenir une pression suffisante sur la membrane. Le transfert est alors réalisé durant 1 heure à 80 V.

La membrane est ensuite récupérée puis recouverte d'un agent chimique (mélange de Mehtylimidazole et de 1-Ethyl-3-(3-diméthylaminopropyl)carbodiimide) pendant 1 heure à 60°C permettant de réaliser un crosslink des ARNs sur la membrane.

7.4 Marquage de la sonde oligonucléotide, Hybridation et Révélation

250 à 300 ng de sonde oligonucléotide est radio marquée par phosphorylation de l'extrémité 5' à l'aide d'une polynucléotide Kinase et en présence de 50 µCi ³²P-α-ATP conformément au protocole du fournisseur. La membrane est ensuite pré hybridée pendant 15 min à 42°C dans le tampon d'hybridation (ULTRAhyb®-Oligo buffer ; Life technologies). La totalité de la sonde radio marquée est ensuite ajoutée directement dans le tampon de pré hybridation et incubée sur la nuit à 42°C.

La membrane est ensuite rincée plusieurs fois avec du 2X SSC/0.5% SDS à 42°C ainsi qu'avec du 1XSSC/0.5% SDS. Un film photographique est ensuite exposé avec la membrane pendant plusieurs heures (en fonction de l'intensité du signal) puis révélé à l'aide d'un Phosphor Imager.

DISCUSSION ET PERSPECTIVES

1. **DUX4, un lien commun unifiant les patients FSHD1 et FSHD2**

Les contractions *D4Z4* seraient à l'origine d'une hypométhylation de la région aboutissant à la levée de la répression transcriptionnelle de *DUX4-FL*. Or l'existence des patients FSHD 2, chez lesquels les contractions *D4Z4* sont absentes, suggèrent un mécanisme alternatif provoquant la levée de la répression de *DUX4*. Néanmoins, ces patients présentent également une hypométhylation de la région *D4Z4* en lien avec la mutation de la protéine SMCHD1 retrouvée chez 79% des patients FSHD2 (Lemmers et al., 2012). *DUX4-FL* est donc également retrouvé chez les patients FSHD2. De plus, il est nécessaire d'avoir l'haplotype 4qA et donc de posséder un signal de polyadénylation indispensable à la stabilisation du transcrit *DUX4*, chez les patients FSHD 1 et FSHD 2. L'expression de *DUX4* constitue donc un lien unifiant les deux groupes de patients et fait de *DUX4* l'un des candidats majeurs susceptible d'être impliqué dans la physiopathologie de la FSHD en générale.

2. **Une fonction pour les différentes isoformes de *DUX4* ?**

2.1 **Récapitulatif des différentes isoformes de *DUX4***

Chaque répétition *D4Z4* de 3,3 Kb renferme une phase ouverte de lecture de 1245 pb suivie d'une région 3'UTR contenant deux introns et un exon facultatif. Néanmoins, dans un contexte 4qA, la toute dernière répétition *D4Z4* (la plus proche des télomères) possède un troisième exon contenant un signal de poly(A) suivi de 4 exons supplémentaires. Plusieurs isoformes de *DUX4* ont ainsi été identifiées selon les tissus, correspondant à différents épissages alternatifs, mais leurs fonctions restent à être élucidées (**Figure 37**).

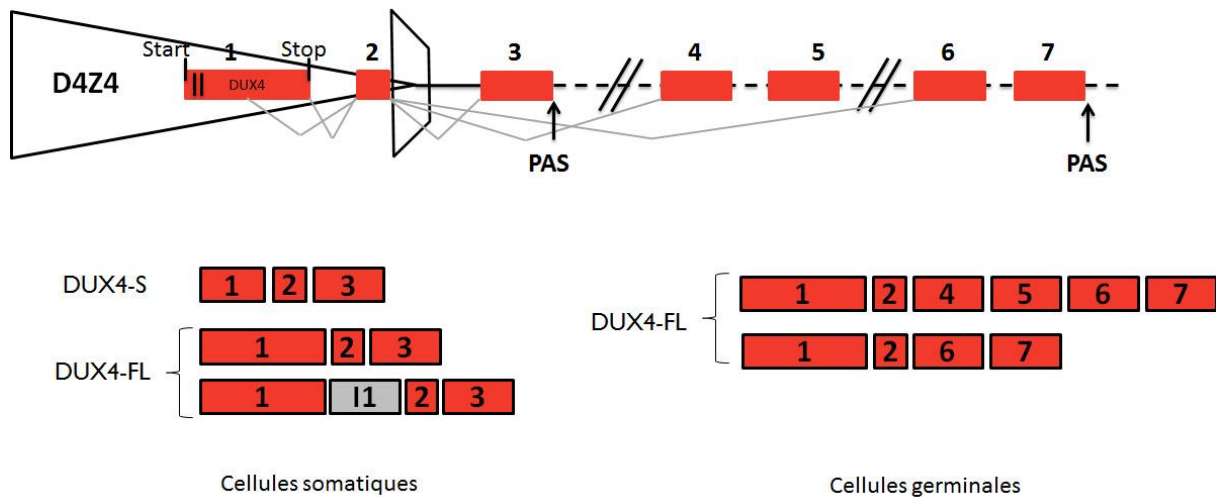


Figure 37 : Schéma de la dernière répétition *D4Z4* associée à l'haplotype 4qA, sur le chromosome 4 et des sites d'épissages alternatifs.

2.2 La production des isoformes *DUX4-S* et *DUX4-FL* régulée par le statut épigénétique de la cellule ?

DUX4 est soumis à un épissage alternatif entre les patients FSHD et Contrôles. Chez les individus contrôles, un événement inconnu provoque l'utilisation d'un site cryptique d'épissage, localisé dans la phase codante de *DUX4*, tronquant la protéine de sa région C-terminale (portant le domaine transactivateur). Néanmoins, il est fortement probable que ce changement d'épissage soit relié au contexte épigénétique de la région. En effet, il a été montré que le changement d'état de méthylation des histones influait sur l'épissage (Luco et al., 2010). Puisque les patients FSHD présentent une hypométhylation de la région *D4Z4*, il est possible d'envisager que cela soit à l'origine de la transition entre les isoformes. Afin de montrer cela, une étude a utilisé des fibroblastes issus d'individus contrôles ou FSHD reprogrammés en cellules souches pluripotentes induites (ips) (Snider et al., 2010). Dans ce système, les fibroblastes contrôles expriment la forme *DUX4-S* associée à une forte méthylation H3K9me3. Une fois reprogrammé en iPS, c'est la forme *DUX4-FL* qui est détectée et associée à une faible méthylation. Enfin, la différenciation des iPS en corps embryonnaire induit une ré-augmentation du niveau de méthylation et un retour de l'isoforme *DUX4-S* (Figure 38).

	<i>DUX4-S</i>	<i>DUX4-FL</i>	Méthylation H3K9me3
Fibroblastes Ctrl	X	-	Forte
iPS Ctrl	-	X	Basse
EB ctrl	X	-	Forte
Fibroblastes FSHD	-	X	Basse
iPS FSHD	-	X	Basse
EB ctrl	-	X	Basse

Figure 38 : Récapitulatif des formes de *DUX4* exprimé et de l'état de méthylation H3K9 dans les différents lignages.

Ces résultats semblent donc montrer que la transition entre l'isoforme courte et longue de *DUX4* serait régulée au moins en partie par l'état de méthylation de la région *D4Z4*.

2.3 Des isoformes alternatives de *DUX4* présentes dans le muscle squelettique des patients FSHD

Contrairement à la forme *DUX4-S* où l'épissage a lieu dans la séquence codante, d'autres isoformes où l'épissage alternatif est localisé dans la région 3'UTR (ne modifiant donc pas la séquence protéique finale) ont été identifiées. Le premier intron est épissé de manière alternative tandis que le second l'est de manière constitutive, générant deux isoformes chez les individus FSHD (*DUX4-FL 1* et *DUX4-FL 2*). L'isoforme *DUX4-FL2*, où le premier intron est retenu, est préférentiellement amplifiée dans les biopsies fœtales FSHD tandis que les 2 isoformes de *DUX4-FL* (épissage ou non de l'intron 1) sont retrouvées chez l'adulte. Il est alors possible d'imaginer que la forme *DUX4-FL2* est préférentiellement produite au cours du développement et que progressivement les deux formes soient co-exprimées. Comme démontré précédemment, l'état de méthylation semble influencer sur l'épissage alternatif et pourrait donc participer à la sélection des isoformes longues.

Afin de déterminer si les deux formes *DUX4-FL* sont produites en égale quantité, nous avons réalisé des PCR à l'aide d'un couple d'amorces encadrant les deux introns. Sur des cellules fœtales FSHD, nous avons été capables d'amplifier simultanément les deux isoformes de *DUX4-FL* mais il semblerait que *DUX4-FL1* soit moins abondant. Nous ne sommes néanmoins pas en mesure de déterminer si les deux formes sont co-exprimées au sein de la

même cellule. Pour cela il serait nécessaire de réaliser des RNA-FISH en utilisant une sonde sur l'intron 1 (détection de *DUX4-FL2*) et une sonde à la jonction des exons 1 et 2 ou l'intron 1 est épissé. De plus, il serait intéressant de déterminer par PCR si *in vitro*, l'expression des différentes isoformes varient en fonction de la différenciation cellulaire.

2.4 DUX4 : un facteur de transcription nécessaire dans les gonades mais toxique dans le muscle ?

Afin de déterminer la fonction d'une protéine, l'analyse de gènes homologues chez d'autres espèces peut fournir des informations sur sa fonction chez l'Homme. Des analyses phylogéniques des répétitions *D4Z4* ont mis en évidence la présence de paralogues chez les primates (Chimpanzee, orangutan), des mammifères africains (familles des Eléphants) ainsi que chez les rongeurs, où l'ORF *DUX4* est conservé depuis environ 100 millions d'années, suggérant un rôle ancestral de *DUX4* (Clapp et al., 2007). Chez la souris, *Duxbl*, paralogue de *DUX4*, est spécifiquement exprimé durant le développement des gonades, dans les oocytes, les spermatogonies et durant les phases précoces de développement du muscle squelettique (Wu et al., 2010) suggérant que *Duxbl* joue un rôle majeur durant la myogenèse et la reproduction. Compte tenu du rôle de *Duxbl*, une équipe s'est intéressée, chez l'Homme, à l'expression de *DUX4* dans des testicules d'individus contrôles (Snider et al., 2010). Il a ainsi été mis en évidence, par PCR et immuno-histochimie, que *DUX4-FL* était présent en dehors d'un génotype pathologique. De plus, il a été montré que la sur-expression de *DUX4* induisait des dérégulations géniques majeures, notamment l'expression de gènes ayant un rôle dans la lignée germinale (Geng et al., 2012). Ces résultats suggèrent donc que *DUX4* pourrait avoir un rôle normal dans les gonades. Chez les individus FSHD, l'absence de répression somatique de *DUX4* dans le muscle squelettique, serait à l'origine d'un gain de fonction toxique médié par l'activation d'un programme gamétogénique non compatible avec le tissu (Geng et al., 2012).

Il faut cependant noter que les isoformes produites dans les gonades ne sont pas similaires à celles observées dans le muscle (Snider et al., 2010). *DUX4* est constitué de deux exons contenus dans chaque répétition *D4Z4* mais la dernière répétition possède une région 3'UTR renfermant 5 exons supplémentaires (exon 3 à 7) présents uniquement sur l'haplotype 4qA. Une étude a montré que dans les gonades, l'exon 3 contenant le signal de poly(A) est épissé pour rejoindre les exons 4 à 7 contenant un second signal de poly(A) localisé dans

l'exon 7 (Snider et al., 2010). De plus, ces transcrits sont issus à la fois du chromosome 4 et du chromosome 10 (Snider et al., 2010), tandis que dans le muscle squelettique de patients FSHD et Contrôles, seul le chromosome 4 serait à même de produire des transcrits DUX4. Il semble donc qu'il existe un mécanisme tissu spécifique bloquant la transcription à partir du chromosome 10 dans le muscle squelettique. Puisque le statut épigénétique de la région semble influencer sur les isoformes de DUX4, il est possible d'envisager qu'une hypométhylation des répétitions *D4Z4* du chromosome 10 provoque une dé-repression transcriptionnelle de DUX4 dans les testicules. Si cette hypothèse est valide, alors chez les individus FSHD 2, où les répétitions *D4Z4* sont hypométhylés à la fois sur les chromosomes 4 et 10 (chez les individus FSHD 1 l'hypométhylation n'est retrouvé que sur le chromosome 4), DUX4 pourrait être produit à partir de ces deux chromosomes.

La variation de l'épissage dans la région 3'UTR n'influençant pas la séquence protéique, nous ne pouvons que spéculer sur sa fonction. Il est possible que la présence d'une région 3'UTR agrandie assure une meilleure stabilité du transcrit ou permette de bloquer les fonctions toxiques de DUX4 grâce à des interactions ARN-Protéine spécifiquement dans les gonades. Pour tester l'hypothèse de la stabilité, il est possible de générer des constructions où la région génomique (exon 1 à 7) a été clonée et où des mutations des sites donneurs et accepteurs d'épissages ont été introduites. L'établissement de lignées cellulaires exprimant stablement ces constructions peuvent ensuite être traité avec un agent bloquant la transcription (Actinomycine D) afin de déterminer la demi vie de l'ARN d'intérêt (quantification par PCR en temps réel).

Il est néanmoins important de souligner qu'à ce jour, aucune étude n'a été publiée confirmant ou infirmant la présence d'isoformes de DUX4 différentes dans les testicules.

2.5 DUX4, un gène, plusieurs protéines ?

L'analyse de la séquence génomique a mis en évidence la présence de trois codons initiateurs pouvant donc permettre la traduction de trois protéines (**Figure 39**). La sur-expression de DUX4 a permis de confirmer par western blot, la production de protéines issues des codons MKG et MQG. Puisque l'anticorps utilisé ne peut détecter la protéine provenant du troisième codon MQG (épitope absent), il n'est pas certain qu'une protéine soit produite à partir de ce codon (Snider et al., 2009, 2010). Il est important de noter que ces résultats ont été obtenus dans des essais de sur-expression, or à l'heure actuelle, il n'existe aucune donnée sur

les formes endogènes de DUX4, dans le muscle squelettique. Les seules détections de DUX4 endogènes ont été réalisées sur des extraits protéiques de testicules (Snider et al., 2010), ou, par western blot, seule une isoforme a été mise en évidence, de plus, la taille détectée n'apparaît pas similaire à celle observée en sur-expression. Plusieurs hypothèses sont susceptibles d'expliquer ces résultats. Tout d'abord, puisqu'il n'existe aucune donnée *in vivo* sur les formes de DUX4 produites dans le muscle squelettique pathologique, il n'est pas possible de déterminer qu'elle forme est produite. Il est également important de rappeler que dans les testicules, les transcrits DUX4 semblent provenir à la fois du chromosome 4 et du chromosome 10 (Snider et al., 2010), par conséquent, il est impossible de savoir qu'elle est l'origine de la protéine dans les testicules.

Ces résultats suggèrent donc que plusieurs protéines peuvent être produites à partir de start alternatif, mais ces isoformes protéiques n'ont à ce jour jamais été identifiées dans des conditions d'expression endogène.

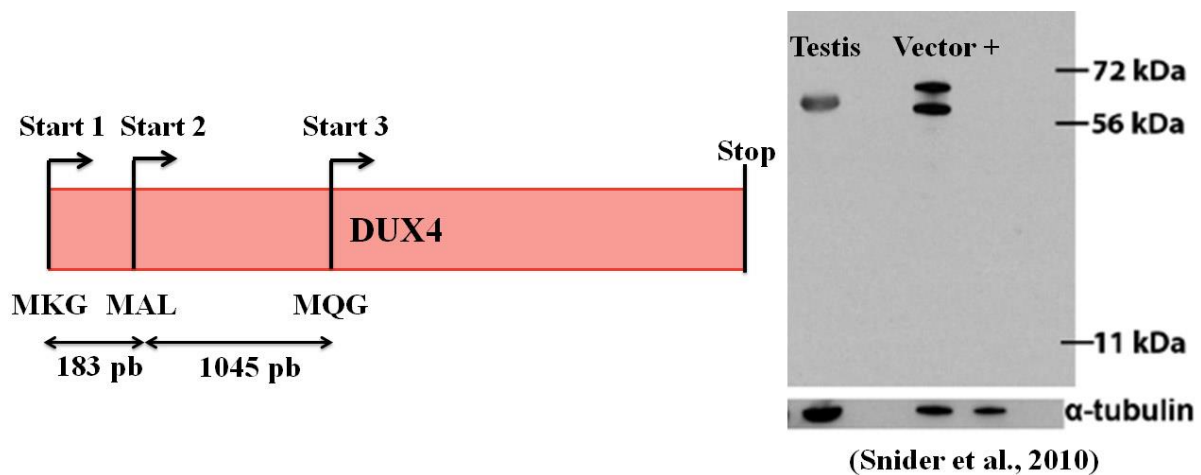


Figure 39 : Start alternatif et codon initiateur dans l'ORF de DUX4 pouvant générer plusieurs isoformes de DUX4 visibles par sur-expression en western blot.

3. Expression de DUX4 dans des cellules et biopsies FSHD

3.1 Cinétique d'expression *in vitro* dans des cellules fœtales FSHD

Dans un premier temps, l'analyse du niveau d'expression de *DUX4* a été réalisée sur des cellules issues d'un quadriceps et d'un trapèze fœtal FSHD1 (4,4 répétitions *D4Z4*). Nous avons ainsi été capables de détecter la présence de DUX4 à la fois au niveau transcriptionnel et au niveau protéique. La quantification du nombre de noyaux DUX4 positif (8,9%) est en accord avec les études publiées et confirme le faible nombre de noyaux exprimant DUX4.

Nous avons mis en évidence que *DUX4-FL* était progressivement exprimé au cours de la différenciation avec un pic d'expression situé entre J4 et J5 suivi d'une diminution. Or nous avons constaté que sur le quadriceps FSHD, le niveau d'expression de *DUX4-FL* semblait ré-augmenter entre J7 et J10 de différenciation pouvant suggérer que DUX4 est exprimé selon un modèle ondulatoire. Compte tenu de la difficulté à maintenir des myotubes en culture pendant une longue période (décollement progressif) nous n'avons pas regardé le niveau d'expression de DUX4 à des stades plus tardifs (15 ou 20 jours de différenciation par exemple) et nous ne pouvons donc pas déterminer si DUX4 est exprimé par vagues successives.

Nous avons ensuite comparé, par PCR quantitative, le niveau d'expression de *DUX4-FL* entre le quadriceps et le trapèze fœtal FSHD et avons mis en évidence sa plus forte expression dans le quadriceps (muscle tardivement atteint chez l'adulte) comparé au trapèze (muscle précocement atteint chez l'adulte). Ces résultats semblent montrer que le niveau d'expression de DUX4 ne serait pas relié à la sévérité d'atteinte future de ces deux muscles. Il est néanmoins important de moduler cette conclusion, en effet, ces résultats ne concernent qu'un seul fœtus FSHD et seule l'augmentation du nombre d'échantillon serait à même de déterminer si la différence d'expression observée est une constante entre ces muscles ou une spécificité intra-individus. Enfin, les atteintes musculaires associées à la FSHD n'étant pas présentes chez les fœtus FSHD, le prélèvement des biopsies est donc fait de manière aléatoire et ne tient donc pas compte de l'asymétrie future. C'est pourquoi il est possible d'envisager que la différence d'expression de DUX4 entre le quadriceps et le trapèze fœtal soit liée au fait que le trapèze prélevé n'aurait pas (ou peu) été cliniquement atteint lors du développement de la pathologie contrairement au quadriceps.

3.2 DUX4 est également détecté *in vivo* dans des biopsies fœtales et adultes FSHD

Bien que la culture cellulaire offre un bon modèle d'étude, il n'est pas exclu que les conditions de cultures puissent moduler le niveau d'expression génique (facteurs de croissance présents dans les milieux de culture...). Afin de s'absoudre de ce biais technique, dans un travail collaboratif, nous avons effectué des analyses d'expression de DUX4 directement sur des ARN extraits de biopsies fœtales FSHD et Contrôles. Nous avons ainsi mis en évidence que DUX4-FL était présent dans 4 quadriceps fœtaux FSHD sur 5. De plus, sa fréquence de détection est fortement réduite dans les biopsies contrôles (détecté dans 5 échantillons sur 21) en accord avec ce qui a déjà été décrit chez l'adulte (Jones et al., 2012). L'analyse de biopsies adultes FSHD et contrôles a abouti aux mêmes résultats, une détection de DUX4FL à une fréquence moins élevée chez les individus contrôles.

Comme décrit, *DUX4-FL* est également détecté dans d'autres tissus notamment dans la peau, le cerveau, le foie ou encore le diaphragme et cela chez des fœtus porteurs ou non de contractions *D4Z4*. Pourtant, ces tissus ne sont pas atteints chez les patients suggérant que, soit un ou plusieurs partenaires de DUX4 sont absents dans ces tissus, soit que finalement DUX4 n'a pas de rôle direct dans l'atteinte musculaire. La seconde hypothèse est probablement vraie car une souris exprimant DUX4 n'a développé aucun symptôme musculaire (Krom et al., 2013), de plus, nos analyses montrent que DUX4 est plus fortement exprimé, *in vitro*, dans un muscle tardivement atteint (quadriceps versus trapèze).

L'expression de DUX4 serait donc un évènement secondaire à la relaxation de la chromatine mais n'ayant potentiellement pas de lien (direct ou indirect) sur le phénotype musculaire.

3.3 Une détection de DUX4 plus difficile dans les biopsies

Contrairement aux cultures cellulaires, l'amplification de *DUX4-FL* n'est possible que par PCR nichée (deux PCR successives ou le produit de la première sert de matrice pour la seconde). La quantification du niveau d'expression de *DUX4-FL* n'est donc pas possible dans ces conditions.

La plus grande difficulté à amplifier *DUX4-FL* dans les biopsies peut avoir plusieurs explications. Tout d'abord, contrairement aux cultures cellulaires *in vitro* constituées dans

l'immense majorité de myoblastes, les biopsies sont elles très hétérogènes (tissus adipeux, fibreux...). Cette hétérogénéité pourrait alors entraîner une dilution du signal pouvant expliquer la difficulté à amplifier DUX4 dans des biopsies.

Enfin, nous pouvons également supposer que l'expression de *DUX4-FL* est artificiellement augmentée (facilitant sa détection) dans les cultures *in vitro*, liées aux conditions de cultures. L'influence des milieux de cultures sur le niveau d'expression de DUX4 a déjà été mise en évidence, en effet, l'utilisation d'un milieu de différenciation particulier augmente de six à huit fois le nombre de noyaux exprimant DUX4 (Block et al., 2013).

4. L'expression de DUX4 est-elle liée aux vagues de myogenèse ?

L'analyse d'expression de DUX4 a été faite à la fois sur des cinétiques de cultures cellulaires et sur des biopsies. Nos observations *in vitro* montrent que *DUX4-FL* est augmenté au cours de la différenciation musculaire, suggérant que *in vivo*, DUX4 serait up régulé durant la myogenèse. Il est alors possible d'imaginer que *in vivo*, DUX4 soit fortement exprimé lors des étapes de myogenèse précoce (lors de la formation du muscle à partir des somites) puis, comme observé *in vitro*, diminue progressivement au cours de la maturation. Il est néanmoins important de faire la différence entre un muscle squelettique déjà formé et des cultures cellulaires où le degré de maturation et l'environnement n'est pas le même (pas innervation...). En effet, la formation des fibres musculaires *in vivo* chez les vertébrés est un processus multiphasique qui implique deux vagues de différenciations successives. La myogenèse primaire a lieu entre la 8^{ème} et la 10^{ème} semaine de développement et produit majoritairement des fibres lentes (caractérisées par la présence de chaînes lourdes de Myosine de type I). La myogenèse secondaire a lieu plus tardivement, vers la 20^{ème} semaine de développement, et est à l'origine de la formation des fibres rapides (Myosine de type II) (Barbet et al., 1991; Butler-Browne et al., 1990). En culture cellulaire, DUX4 est exprimé au cours de la différenciation mais *in vivo*, lorsque le muscle est déjà formé, il n'existe plus de myogenèse excepté lorsque le muscle est lésé (déchirure musculaire...).

Si l'expression de DUX4 est stimulée lors de la myogenèse, une fois le muscle formé, seules des lésions musculaires, activant la régénération à partir des cellules satellites, seraient

en mesure de provoquer un nouveau cycle d'expression de DUX4 (**Figure 40**). Dans ce cas de figure, l'accumulation de traumatismes musculaires au cours de la croissance pourrait participer à l'effet néfaste de DUX4. Ce modèle serait susceptible d'expliquer pourquoi, alors que les événements moléculaires de la FSHD sont retrouvés dès le stade fœtal, la pathologie ne se déclare que plus tardivement au cours de l'adolescence.

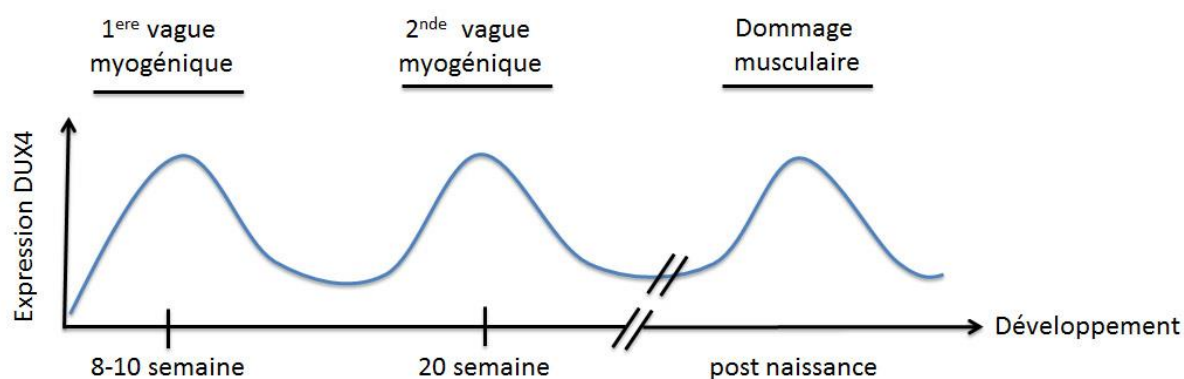


Figure 40 : Modèle hypothétique d'expression de DUX4 en lien avec les vagues myogéniques lors du développement.

5. Dérégulations de gènes impliqués dans la myogenèse

La FSHD étant une pathologie d'origine musculaire, des analyses histologiques sur des coupes musculaires adultes ont été réalisées, mais n'ont pas mis en évidence de défauts structuraux majeurs hormis des modifications de la taille des fibres, des infiltrations adipocytaires, la présence de tissus fibrotiques ou encore une réorganisation du sarcolemme (Lassche et al., 2013; Reed et al., 2006). L'essentiel des défauts observés sont des signes classiques de dystrophie. La recherche de gènes candidats ayant un rôle dans le muscle et induisant les phénotypes musculaires à ce stade n'est pas simple. En effet, il n'est pas possible de déterminer si les défauts géniques observés ont provoqué ou sont provoqués par l'altération du muscle squelettique, de plus il est probable que les deux événements s'entretiennent mutuellement.

C'est pourquoi l'analyse de l'expression de facteurs myogéniques au cours du développement, et donc en absence de toute altération musculaire, s'avère très informative. Dans un travail collaboratif avec l'équipe de Frédérique Magdinier, la comparaison de biopsies fœtales provenant de 6 fœtus FSHD 1 avec 21 fœtus ne portant pas de contractions a

mis en évidence la dérégulation d'un grand nombre de gènes impliqués dans le programme myogénique. Le sarcoglycan, la calpain 3 ou encore la dysferlin (des protéines structurales du muscle squelettique) ont été identifiés comme sur-exprimés spécifiquement chez les fœtus FSHD. D'autres gènes codant des protéines impliquées dans le remodelage du muscle (MURF1, TRIM63), la voie du calcium (RYR1) ou encore la différenciation musculaire (MyoD, MRF4), sont retrouvés sur-exprimés. Les altérations d'expression de gènes impliqués dans la myogenèse sont classiquement associées à des défauts musculaires. Or pour la première fois, cette étude a mis en évidence qu'au stade fœtal, avant toute apparition de signes dystrophiques, une cascade de gènes impliqués dans la myogenèse est dérégulée. L'altération des facteurs myogéniques est en accord avec une étude qui a relié la sur-expression de DUX4 à une augmentation de l'expression de MyoD (Sharma et al., 2013). De même, des résultats préliminaires dans lesquels nous avons réalisé des immunomarquages de MyoD sur nos cultures cellulaires fœtales issues d'un quadriceps et d'un trapèze FSHD confirment que *in vitro*, MyoD est plus fortement exprimé dans le quadriceps FSHD comparé au quadriceps contrôle (cellules provenant d'un fœtus du même âge de développement) (**Figure 41**). De manière intéressante, MyoD est plus fortement exprimé dans les cellules où DUX4 est le plus fort (quadriceps) comparé au trapèze pouvant suggérer un lien entre les deux événements. Afin de déterminer s'il existe un lien entre le niveau d'expression de MyoD et DUX4 il serait nécessaire de réaliser les mêmes analyses sur les cellules traitées ou non avec le ShARN DUX4.

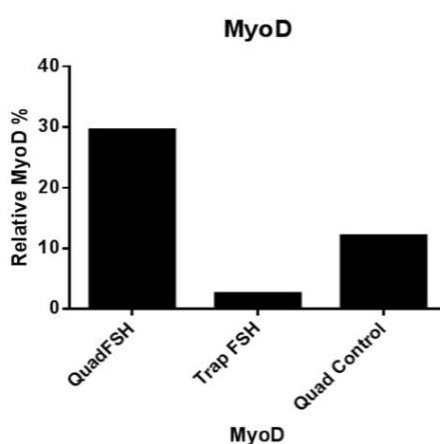


Figure 41 : Quantification du nombre de noyaux MyoD positifs par immunofluorescence sur des cellules fœtales FSHD et contrôle au cinquième jour de différenciation.

Il est néanmoins important de noter que d'autres études où DUX4 est sur-exprimé provoquaient la diminution de MyoD (Bosnakovski et al., 2008; Snider et al., 2009) et une étude protéique a également mis en évidence une moins forte expression de MyoD dans des biopsies FSHD (Celegato et al., 2006). L'ensemble de ces résultats contradictoires (sur-expression ou sous-expression de MyoD) montre néanmoins un point commun, la dérégulation de facteurs myogéniques qui pourrait être médiée par DUX4. Il est alors possible d'envisager que l'apparition de défauts dans l'expression de facteurs myogéniques à des stades précoces de développement, provoque des défauts structurels du muscle participant, en présence de DUX4, à la physiopathologie de la FSHD.

Des études visant à déterminer s'il existait des altérations de la différenciation ou de la fusion, ont alors été réalisées sur des cultures cellulaires de patients FSHD. Une fois de plus les résultats obtenus entre différentes équipes sont contradictoires, certains montrant des altérations dans le processus de fusion (Barro et al., 2010) et d'autres ne trouvant aucune différence (Krom et al., 2012; Tsumagari et al., 2011). Néanmoins, des tests préliminaires que nous avons réalisés sur deux cultures de cellules fœtales exprimant ou non un ShARN dirigé contre DUX4 semblent montrer que l'extinction de DUX4 serait associée à une fusion améliorée. Ces résultats intéressants doivent néanmoins être répétés et confirmés sur d'autres cultures cellulaires.

Ainsi bien que la FSHD soit considérée comme une dystrophie musculaire de l'adulte, ces résultats suggèrent plutôt que les mécanismes moléculaires prennent naissance dès le développement précoce du muscle pouvant induire une prédisposition de certains groupes musculaires.

6. Altération de l'expression génique en cis et en trans chez le fœtus porteur de contractions *D4Z4*

Bien que les modèles de sur-expression de DUX4 induisent un grand nombre de dérégulations géniques, plusieurs autres modèles, notamment basés sur des changements conformationnels de la chromatine, ont été mis en avant afin d'expliquer le rôle des répétitions *D4Z4* et l'impact des contractions sur l'expression de gènes du locus 4q35 (par effets de positions). Dans ces modèles, la relaxation de la chromatine associée à la perte de marques répressives (méthylation des résidus H3K9) (Zeng et al., 2009), la perte de domaines d'attache nucléaire liée aux contractions (Petrov et al., 2006) ou encore l'interaction de CTCF

et des lamines de type A avec les répétitions D4Z4 chez les patients FSHD 1 et FSHD 2 (Ottaviani et al., 2009) serait à l'origine de dérégulations géniques.

Néanmoins, la plupart des gènes décrits comme dérégulés sont controversés et les résultats sont dépendants du matériel biologique utilisé (biopsies, myoblastes ou myotubes primaires) ainsi que des techniques employées (surexpression de DUX4 ne mimant pas le niveau endogène extrêmement bas).

6.1 Gènes du locus 4q35

Dans un travail collaboratif avec l'équipe de Frédérique Magdinier, nous avons alors entrepris l'analyse de plusieurs gènes du locus 4q35 sur des biopsies fœtales FSHD afin de déterminer si l'expression des gènes de ce locus était altérée par la présence des contractions D4Z4 au stade fœtal. Parmi 36 gènes étudiés du locus 4q35, 14 apparaissent comme sur-exprimés spécifiquement dans les biopsies fœtales FSHD.

Parmi eux figure *FRG1*, un gène déjà décrit comme sur-exprimé dans des biopsies adulte FSHD (Gabellini et al., 2002). L'expression de *FRG1* a été montrée comme importante pour le développement du muscle squelettique dans un modèle de *Xenopus laevis* (Hanel et al., 2009) et une sur-expression de *FRG1* dans ce modèle aboutit à une formation anormale des muscles Epaxial (muscles du dos) et Hypaxial (muscles des membres et du tronc). *FRG1* joue également un rôle dans la maturation des préARNm (épissage) et il a été montré que sa sur-expression chez la souris induit des altérations de l'épissage de différents transcrits. Puisque l'une des isoformes de *DUX4-FL* est préférentiellement retrouvée dans les fœtus FSHD, chez lesquels *FRG1* est retrouvé sur-exprimé (Broucqsault et al., 2013), il est possible d'envisager que la détermination des transcrits *DUX4-FL* est dépendante du niveau d'expression de *FRG1*. Néanmoins, le rôle des isoformes de *DUX4-FL* n'est à ce jour pas élucidé.

Nous avons également mis en évidence que le gène *ANT1*, codant l'adénine nucléotide translocator 1, est sur-exprimé spécifiquement dans les biopsies fœtales FSHD. Ce gène a également été identifié comme sur-exprimé mais dans des biopsies adultes FSHD (Gabellini et al., 2002; Laoudj-Chenivresse et al., 2005) néanmoins, tout comme *FRG1*, il n'a pas été confirmé par d'autres laboratoires (Jiang et al., 2003; Osborne et al., 2007) et nos analyses sur des biopsies adultes n'ont pas mis en évidence cette dérégulation. ANT1 est une protéine membranaire jouant un rôle dans le transport de l'ADP de la matrice mitochondriale vers le

cytoplasme. Il a été caractérisé que les cellules de patients FSHD présentent une plus grande sensibilité aux espèces réactives de l'oxygène (ROS) (Barro et al., 2010; Turki et al., 2012; Winokur et al., 2003) associé à des sur-expressions d'agents anti-oxydants comme la superoxyde dismutase ou la catalase (Bauer et al., 1999; Laoudj-Chenivresse et al., 2005). On parle de stress oxydant lorsqu'il y a un déséquilibre entre agents oxydants et anti-oxydants et une altération dans le mécanisme de transport des électrons permettant la synthèse d'énergie est l'une des causes d'apparition d'un stress oxydant.

Il est possible d'envisager que la sur-expression d'ANT1 au stade fœtal soit à l'origine d'une plus grande activité de la chaîne respiratoire conduisant à une plus forte production de ROS dans les cellules et donc à une plus forte sensibilité au stress oxydant. Néanmoins, en culture, il a été montré que la sensibilité des cellules FSHD au stress oxydant n'était pas différente des cellules contrôles au stade myotubes (Winokur et al., 2003). Par conséquent, si la sur-expression d'ANT1 participe à un effet toxique via la production de ROS, cet effet ne serait que transitoire (à des stades myoblastes précoces lors du développement) car une fois la fibre musculaire formée, la sensibilité au stress oxydant est diminuée. De même, des études ont associé un stress oxydant directement lié à la sur-expression de DUX4 (Bosnakovski et al., 2008a). Puisque DUX4 est exprimé dès le stade fœtal, nous pouvons émettre l'hypothèse d'un effet cumulatif entre la sur-expression d'ANT1, DUX4 et un stress oxydant. Il serait donc intéressant de déterminer si les cultures fœtales FSHD présentent une sensibilité au stress oxydant augmenté, associé ou non à la sur-expression d'ANT1, et si ce stress varie en fonction de la présence de DUX4 (utilisation de ShARN dirigé contre DUX4).

D'autres gènes controversés du locus 4q35 ont également été analysés comme *FRG2* ou encore *ALP* (intervient dans l'assemblage du cytosquelette au niveau des stries Z du muscle squelettique) mais aucune différence d'expression n'a été identifiée dans les biopsies fœtales et adultes porteuses des contractions *D4Z4*.

De manière intéressante, un certain nombre des gènes dérégulés sont localisés à environ 5Mb de la première répétition *D4Z4*, dans une région chromosomique décrite comme étant en périphérie nucléaire (Ottaviani et al., 2009, 2010). Ainsi, dans un contexte contracté (valable également lorsque la région *D4Z4* est hypométhylé chez les patients FSHD2), la relocalisation de la région sub-télomérique 4q35 en périphérie nucléaire serait à l'origine de la dérégulation de gènes à proximité. Puisque l'essentiel de ces défauts ne sont pas retrouvés dans des biopsies adultes, il est possible d'envisager que l'altération de ces gènes en lien avec DUX4,

provoque des dommages à des stades précoces du développement musculaire à l'origine de la pathologie.

6.2 Dérégulations géniques en lien avec DUX4 chez le fœtus et l'adulte FSHD

L'analyse de l'expression d'un grand nombre de gènes, décrits comme dérégulés par DUX4, a mis en évidence que ceux-ci étaient progressivement sur-exprimés au cours de la différenciation et selon le même modèle d'expression que DUX4. La diminution du niveau d'expression de ces gènes cibles dans des cellules exprimant un ShARN dirigé contre DUX4, confirme que DUX4 est bien à l'origine de ces dérégulations. *In vivo*, la comparaison de 16 biopsies contrôles (quadriceps fœtaux âgé de 14 à 27 semaines de développement) avec 5 quadriceps fœtaux FSHD a permis de ne confirmer que 2 gènes (MBD3L2 et TRIM43) up-régulés spécifiquement chez le fœtus FSHD (sur les 4 confirmés dans les cellules). Ces 2 gènes sont également les seuls retrouvés dérégulés (parmi ceux testés) dans des quadriceps adultes FSHD. De manière intéressante, le niveau de transactivation de ces 2 gènes est similaire entre les biopsies fœtales et adultes FSHD suggérant que le niveau d'expression de DUX4 est lui aussi similaire. Ces deux gènes confirmés *in vitro* puis *in vivo*, ne seraient normalement pas exprimés (ou très faiblement) dans le tissu musculaire notamment car ils auraient un rôle dans les cellules germinales (Geng et al., 2012). C'est pourquoi l'hypothèse d'une activation d'un programme gamétogénique (médié par DUX4) incompatible avec le développement du muscle squelettique a été proposée mais reste à confirmer.

Puisque la pathologie ne se déclare que durant l'adolescence, il apparaît donc que l'expression de DUX4 et la modulation des gènes cibles seules ne suffisent pas au déclenchement de la FSHD.

7. Comment cette protéine faiblement exprimée peut induire des dérégulations géniques majeures ?

Bien que des essais de sur-expression aient montré que DUX4 était capable d'induire des dérégulations géniques majeures (Geng et al., 2012), l'expression très faible de DUX4 (0,5 à 10% des noyaux) de manière endogène est suffisante pour induire des défauts moléculaires. DUX4 étant un facteur de transcription, la protéine retourne dans le noyau grâce

à des séquences de localisation nucléaire (NLS) (Corona et al., 2013) or cette séquence oriente la protéine vers les noyaux à proximité. Le muscle squelettique étant composé de filaments multinuclés (partageant donc un cytoplasme commun) il est possible d'envisager que DUX4 soit transcrit dans un petit nombre de noyaux puis, après traduction dans le cytoplasme, diffuse dans d'autres noyaux propageant ainsi son effet au sein même de noyaux n'exprimant pas le DUX4. Ce phénomène ayant déjà été décrit dans différentes études (Blau et al., 1983; Ralston and Hall, 1989) nous avons utilisé des systèmes de co-culture entre des myoblastes Humain et Murin (formation d'hétérokaryon hybride) et mis en évidence que DUX4 était bien capable de diffuser entre noyaux. Cette information précieuse permettrait d'expliquer comment une protéine si faiblement exprimée peut être à l'origine de dérégulations géniques globales et participer à la diffusion d'un phénomène pathologique (Figure 42).

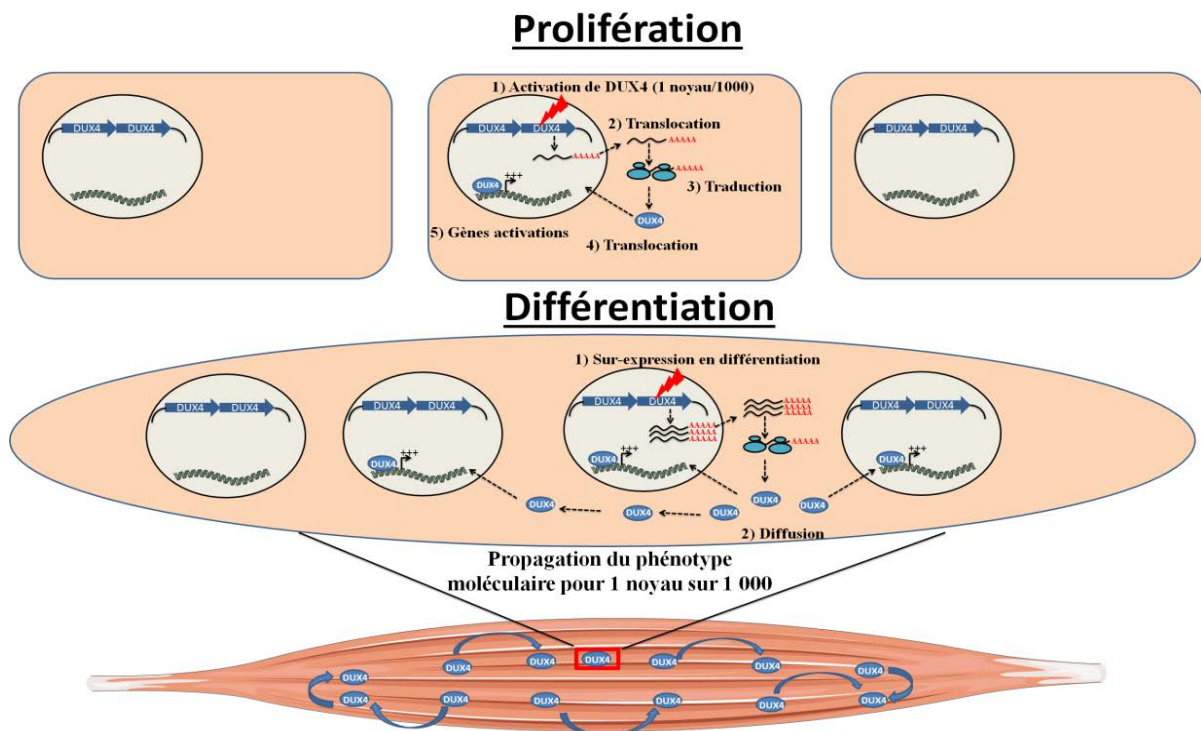


Figure 42 : Proposition du modèle d'expression de DUX4 et de la propagation du phénotype moléculaire dans un muscle entier : En prolifération, DUX4 est exprimé dans un nombre limité de noyaux (1 sur 1000). DUX4 est ensuite progressivement sur-exprimé en différenciation. La proportion de noyaux au sein des myotubes/fibres musculaires transcrivant DUX4 resterait faible mais les différents noyaux étant regroupés au sein d'un même cytoplasme, dans lequel DUX4 est traduit, permettraient une diffusion de la protéine dans les noyaux situés à proximité et assureraient la propagation du phénotype moléculaire dans un muscle entier.

Ce modèle démontre donc qu'il n'est pas nécessaire que la protéine d'intérêt soit produite dans l'ensemble des noyaux. DUX4 est toxique en sur-expression, certainement due aux effets spécifiques et non spécifiques liés à l'expression massive d'un facteur de transcription. Or, dans des conditions d'expression endogène il ne semble pas y avoir de toxicité apparente. Compte tenu du phénomène de diffusion, il est possible d'imaginer que cela induise une « titration » de la protéine liée à sa dilution entre les noyaux situés à proximité, diminuant ainsi son effet toxique et pouvant alors expliquer pourquoi DUX4 apparaît comme moins toxique dans des myotubes en culture cellulaire.

8. Un autre gène retrouvé dérégulé chez le fœtus porteur de contractions *D4Z4* : *FAT1*

L'une des caractéristiques de la FSHD est l'atteinte asymétrique des différents groupes musculaires. Aucun lien n'a pu être établi entre le côté de l'atteinte et le caractère gaucher ou droitier des patients (Brouwer et al., 1992). D'un point de vue moléculaire, il ne semble pas que DUX4 puisse avoir un rôle dans l'asymétrie car il est retrouvé exprimé à la fois dans des muscles atteints et non atteints (Jones et al., 2012), de plus, une souris sur-exprimant DUX4 ne montre pas d'atteinte musculaire (Krom et al., 2013). Il serait néanmoins intéressant de comparer si au stade fœtal, en dehors de toute atteinte musculaire et d'influence sur le caractère gaucher-droitier, un muscle droit et gauche présente le même niveau d'expression de DUX4.

A l'heure actuelle, il a été montré que la sous-expression du gène *FAT1* chez la souris récapitulait la majorité des symptômes musculaires liés à la FSHD, notamment une atrophie asymétrique (Caruso et al., 2013) et cette sous-expression est également retrouvée spécifiquement chez les fœtus porteurs des contractions *D4Z4*. De manière intéressante, la sévérité d'atteinte musculaire est directement reliée à la quantité d'expression de *FAT1* chez la souris, il est donc possible d'envisager que chez les patients FSHD, le niveau endogène de *FAT1* dans les différents groupes musculaires puisse participer à la prédisposition de l'atteinte musculaire et déterminer la cinétique progressive de la pathologie.

8.1 Un lien entre DUX4 et FAT 1 ?

Si FAT1 joue un rôle dans l'atteinte musculaire, agit-il en lien avec DUX4 ou de manière indépendante ? la sur-expression de DUX4 a mis en évidence que FAT1 était sous-exprimé d'un facteur 2 (Geng et al., 2012) ainsi selon cette étude, DUX4 serait capable de réguler négativement l'expression de FAT1. Néanmoins, des analyses que nous avons effectuées au laboratoire n'ont pas confirmé ces résultats (**Figure 43**). En effet, l'utilisation de cellules fœtales exprimant constitutivement un ShARN dirigé contre DUX4, ne montre pas de modification du niveau d'expression de FAT1, suggérant que DUX4 ne joue pas de rôle dans la sous-expression de FAT1 chez les fœtus porteurs de contractions *D4Z4*.

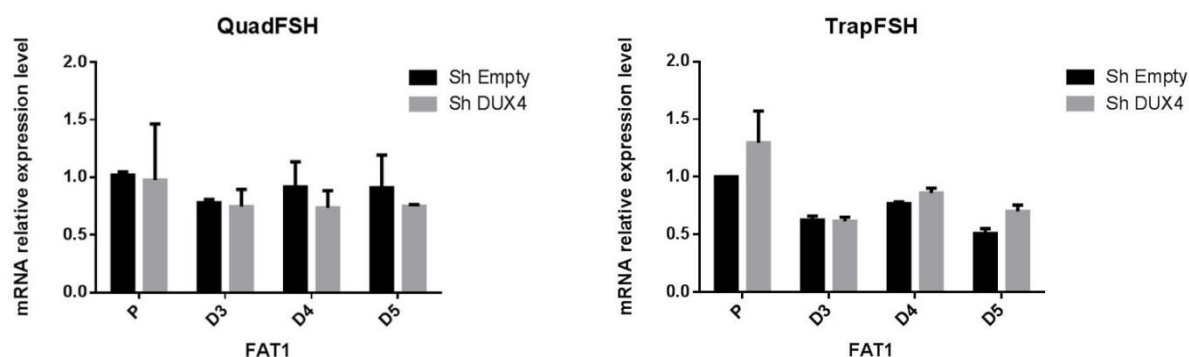


Figure 43 : Quantification du niveau d'expression de FAT1 par RT-qPCR dans les cellules fœtales FSHD exprimant ou non un ShARN dirigé contre DUX4 en différenciation.

Ces résultats sont en opposition avec ceux décrits par l'étude précédente, mais il est probable que l'approche utilisée dans cette étude, consistant à sur-exprimer massivement un facteur de transcription, augmente considérablement les probabilités de faux positifs, contrairement à la stratégie visant à diminuer le niveau d'expression endogène de DUX4.

Ainsi, nos résultats démontrent que la diminution de l'expression de FAT1 n'est pas liée à une régulation négative par DUX4. L'analyse du niveau d'expression de FAT1 dans 3 quadriceps fœtaux porteurs de 1,5 ; 4,3 et 7 répétitions *D4Z4* suggèrent (n=1) que FAT1 est plus fortement sous-exprimé dans les fœtus porteurs d'un faible nombre de répétitions (Caruso et al., 2013). FAT1 étant localisé sur la même portion chromosomique que les répétitions, il est possible d'imaginer un mécanisme épigénétique par lequel les contractions *D4Z4* entraînent une répression épigénétique de FAT1.

9. Quelles origines pour expliquer les atteintes musculaires différentielles ?

9.1 Origine somitique ?

Durant le développement embryonnaire, il y a formation de structures issues du mésoderme nommées somites. On peut séparer les somites en deux parties: les parties ventrales et dorsales. La partie ventrale du somite (le sclérotome) participera à la formation du cartilage et des os. La partie dorsale (dermomyotome) formera le muscle squelettique. Au cours de la maturation, les somites seront influencés par leur environnement qui est différent en fonction de leurs localisations. Les cellules du dermomyotome prolifèrent pour donner d'autres structures: les cellules des myotomes hypaxial et épaxial. Les cellules du myotome hypaxial migrent et donnent les muscles des membres. Les cellules du myotome épaxial forment les muscles du dos et le diaphragme (Braun and Gautel, 2011; Perry and Rudnick, 2000). Les différents somites sont à l'origine des différents groupes musculaires, par exemple, les 7 premiers somitomères vont former les muscles de la face, de la mâchoire et du larynx. Dans la FSHD, seulement 2 groupes musculaires de la face sont touchés (obicularis oculi et l'obicularis oris) tandis que les muscles de la mâchoire ne sont pas atteints. Compte tenu de cela, il n'est pas concevable que l'atteinte musculaire différentielle ait une origine somitique car dans ce cas, l'intégralité des muscles issus de ces somites serait atteinte.

Bien que les cellules du muscle squelettique expriment les 4 mêmes facteurs de transcription au cours du développement comme MYF5 (Myogenic factor5), MRF4 (muscle-specific regulatory factor 4), MYOD (Myoblast determination protein) et la Myogenin (MYOG) (Pipes et al., 2006), les cascades d'activations géniques sont différentes en fonction du type musculaire futur. En effet, dans les muscles du tronc, l'expression des 4 facteurs myogéniques est induite par PAX3 tandis que dans les muscles des membres, elle est induite par PAX3, SIX (sine oculis homeobox homologue) et EYA (eyes absent) (**Figure 44**). Considérant que DUX4 est exprimé dès le stade fœtal et qu'il peut interférer avec les facteurs myogéniques comme PAX3, PAX7 ou MyoD, il est possible d'envisager qu'à ces stades très précoces du développement, DUX4 puisse altérer la mise en place du programme myogénique et que son impact n'est pas le même entre les groupes musculaires. Néanmoins ce modèle ne peut expliquer la présence d'une atteinte musculaire asymétrique qui est l'une des caractéristiques de la FSHD.

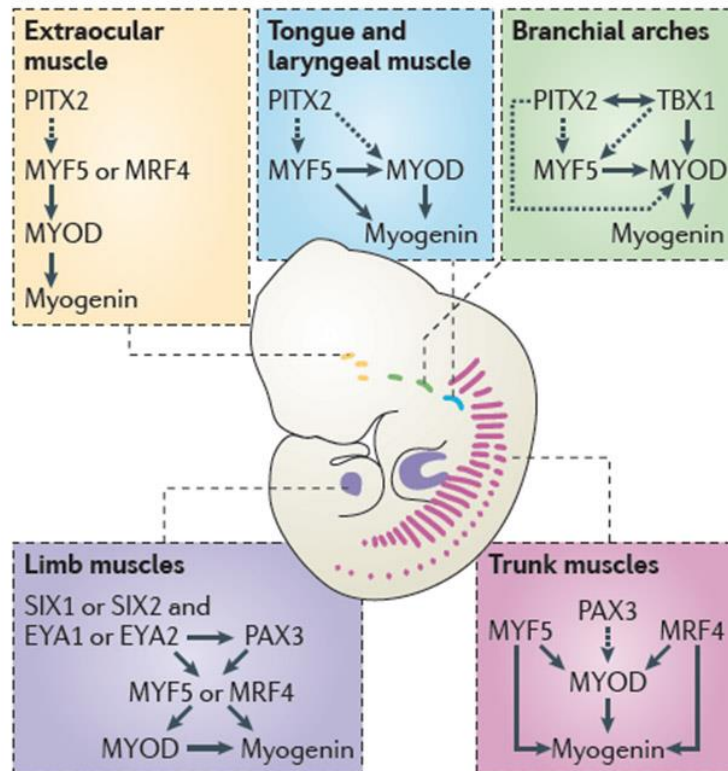


Figure 44 : Voies d'activations du programme myogénique en fonction du type musculaire (Braun et al., 2011).

9.2 Différences intrinsèques entre les muscles ?

Compte tenu des différences développementales entre les groupes musculaires, il est possible d'envisager que des différences intrinsèques entre les muscles les prédisposent à une atteinte précoce. Pour vérifier cela, nous avons dans un premier temps regardé si la cinétique de fusion/différenciation était similaire entre les cellules issues d'un quadriceps et d'un trapèze fœtal FSHD. La quantification de l'expression de la myogénine (facteur de transcription impliqué dans l'activation du programme myogénique) ainsi que de l'inclusion de l'exon 11 de BIN1 (intervenant dans la maturation des tubules T et est donc spécifiquement augmenté au cours de la différenciation musculaire) a mis en évidence un délai de leurs expressions dans le trapèze, suggérant que l'activation du programme myogénique dans ce muscle est retardé. Bien que ce décalage coïncide parfaitement avec la cinétique d'expression de DUX4, les mêmes résultats ont été obtenus sur 8 cultures fœtales contrôles (provenant de 4 fœtus contrôles, 1 quadriceps et 1 trapèze par fœtus). L'absence de DUX4 dans ces cultures contrôles démontre que le retard dans la mise en place du programme

myogénique entre le quadriceps et le trapèze n'est pas lié à l'expression de DUX4 mais semble être une caractéristique intrinsèque du muscle.

Compte tenu des différences dans la mise en place du programme myogénique entre les deux muscles testés, il est possible d'envisager que d'une manière plus générale, dans un contexte pathologique, ces disparités puissent être reliées à l'atteinte plus ou moins précoce de certain groupe musculaire.

10. Finalement, DUX4 a-t-il un rôle dans la pathologie ?

Tout au long de ce manuscrit, un certain nombre d'hypothèses s'affrontent, certaines expliquant comment DUX4 pourrait être en lien avec la pathologie et d'autres non. Le premier point argumentant un rôle de DUX4 est que l'absence de répétition n'entraîne pas la maladie, par conséquent, un élément à l'intérieur des répétitions *D4Z4* (ORF DUX4) semble indispensable au développement de la pathologie. Un contre argument est que dans les répétitions, plusieurs autres domaines, permettant des interactions avec des protéines (enhancer avec KLF15 et insulateur avec CTCF/Lamin), ont été caractérisés. Ces interactions semblent augmenter chez les individus FSHD et participeraient à un changement conformationnel de la chromatine à l'origine des altérations géniques au locus 4q35. L'absence de ces domaines empêcherait le remodelage local de la chromatine, par exemple, la fixation de CTCF associée aux Lamines de type A/C permettrait de rediriger la région sub-télomérique en périphérie nucléaire (Ottaviani et al., 2009) or l'augmentation du nombre de répétitions diminue fortement ces interactions. L'absence de répétitions *D4Z4* aurait donc le même effet que les grandes répétitions, une absence de fixation de CTCF empêchant la relocalisation du sub-télomère. Une seconde hypothèse en faveur d'un rôle de DUX4 est la nécessité de posséder l'haplotype 4qA qui est associé à la présence d'un signal de polyadénylation indispensable à la stabilisation du transcrit DUX4. L'expression de DUX4 est également un point rassemblant les patients FSHD 1 et FSHD 2.

La présence de DUX4 semble donc indispensable mais son rôle précis reste encore à définir. En effet, bien que décrit comme étant une protéine toxique pour le muscle squelettique, cette conclusion ne semble pas être valide à des niveaux d'expression endogène. *In vitro*, les cultures FSHD et Contrôle ne présentent pas de différences phénotypiques notamment au niveau de la fusion ou de la mortalité (pas de différences d'expression de

marqueurs d'apoptose comme p53). L'expression de DUX4 ne semble pas être à l'origine de la sévérité d'atteinte musculaire puisque son niveau d'expression est moindre dans un muscle précocement atteint chez l'adulte, néanmoins, une sensibilité accrue à l'effet de DUX4 dans ces muscles n'est pas à exclure. Enfin, *in vivo*, DUX4 induit (directement ou indirectement) des dérégulations géniques de plusieurs gènes (MBD3L2 et TRIM43) dont le rôle reste à préciser. Il apparaît néanmoins que ces deux gènes ne sont normalement pas exprimés dans le muscle squelettique suggérant que leurs activations contribuent à la pathologie. Il est cependant important de remarquer qu'un modèle de souris exprimant DUX4 n'a développé aucun phénotype musculaire.

Prises ensemble, ces données suggèrent que la FSHD est une pathologie multigénique où les changements de structures chromatinienne liées aux contractions et à l'état de méthylation aboutissent à des dérégulations géniques globales. Associé à cela, le rôle de DUX4 n'est pas à ce jour clair mais il apparaît tout de même indispensable au développement de la pathologie.

CONCLUSION GENERALE

Notre travail a permis de mettre en évidence, pour la première fois, que les caractéristiques moléculaires de la FSHD (expression de DUX4, dérégulation de gènes cibles) sont présentes dès le stade fœtal. L'identification d'un certain nombre de gènes dérégulés, impliqués dans la fonction ou la structure musculaire, suggèrent une altération du programme myogénique à des stades très précoces, pouvant alors prédisposer le muscle à une atteinte future. Malgré le faible niveau d'expression de DUX4, nous avons mis en évidence que DUX4 était capable de propager la cascade de dérégulations qu'il induit, en diffusant d'un noyau à l'autre au sein d'une fibre musculaire, pouvant ainsi expliquer comment une protéine transcrite dans un faible nombre de noyaux peut entraîner des dérégulations géniques majeures.

Contrairement à d'autres pathologies musculaires comme la myopathie de Duchenne ou la mutation du gène codant la dystrophine conduit à une altération de la structure du muscle squelettique, la FSHD n'est pas liée à la mutation d'un gène en particulier mais serait multifactorielle. L'expression de DUX4 seul, dont la toxicité à des niveaux endogènes n'est pas établie, n'apparaît pas suffisante pour induire la pathologie. Les modifications chromatinienne induites par les contractions D4Z4, associées à des changements du statut épigénétique, la dérégulation d'un grand nombre de gènes, en lien ou non avec DUX4, participent ensemble au développement de la pathologie compliquant les approches thérapeutiques.

Néanmoins, bien que notre travail oriente la FSHD comme une pathologie prenant son origine durant le développement précoce, le délai d'apparition des symptômes suppose que les mécanismes aboutissant aux phénotypes pathologiques sont lents et certainement cumulatifs au cours du temps, permettant alors d'envisager des stratégies thérapeutiques chez l'enfant afin de bloquer la progression de la pathologie. L'une des approches envisageable serait de venir bloquer l'expression de DUX4 car, bien que son expression seule n'est pas suffisante, l'existence d'individus sains présentant une délétion de l'intégralité des répétitions D4Z4 (et donc l'absence de DUX4) indique que la présence de DUX4 serait indispensable à la physiopathologie de la dystrophie facio-scapulo-humérale.

BIBLIOGRAPHIES

- Aukerman, M.J., and Sakai, H. (2003). Regulation of flowering time and floral organ identity by a MicroRNA and its APETALA2-like target genes. *Plant Cell* 15, 2730–2741.
- Awerbuch, G.I., and Wishnow, R. (1990). Beevor's sign and facioscapulohumeral dystrophy. *Arch Neurol* 47, 1208–1209.
- Bakker, E., Wijmenga, C., Vossen, R.H., Padberg, G.W., Hewitt, J., van der Wielen, M., Rasmussen, K., and Frants, R.R. (1995). The FSHD-linked locus D4F104S1 (p13E-11) on 4q35 has a homologue on 10qter. *Muscle Nerve Suppl* S39–44.
- Bakker, E., Van der Wielen, M.J., Voorhoeve, E., Ippel, P.F., Padberg, G.W., Frants, R.R., and Wijmenga, C. (1996). Diagnostic, predictive, and prenatal testing for facioscapulohumeral muscular dystrophy: diagnostic approach for sporadic and familial cases. *J Med Genet* 33, 29–35.
- Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* 410, 120–124.
- Barbet, J.P., Thornell, L.E., and Butler-Browne, G.S. (1991). Immunocytochemical characterisation of two generations of fibers during the development of the human quadriceps muscle. *Mech. Dev.* 35, 3–11.
- Barro, M., Carnac, G., Flavier, S., Mercier, J., Vassetzky, Y., and Laoudj-Chenivesse, D. (2010). Myoblasts from affected and non-affected FSHD muscles exhibit morphological différenciation defects. *J Cell Mol Med* 14, 275–289.
- Bartel, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116, 281–297.
- Bartel, D.P. (2009). MicroRNAs: target recognition and regulatory functions. *Cell* 136, 215–233.
- Bastress, K.L., Stajich, J.M., Speer, M.C., and Gilbert, J.R. (2005). The genes encoding for D4Z4 binding proteins HMGB2, YY1, NCL, and MYOD1 are excluded as candidate genes for FSHD1B. *Neuromuscul Disord* 15, 316–320.
- Bauer, M.K., Schubert, A., Rocks, O., and Grimm, S. (1999). Adenine nucleotide translocase-1, a component of the permeability transition pore, can dominantly induce apoptosis. *J. Cell Biol.* 147, 1493–1502.
- Beckmann, J.S., and Spencer, M. (2008). Calpain 3, the “gatekeeper” of proper sarcomere assembly, turnover and maintenance. *Neuromuscul. Disord. NMD* 18, 913–921.
- Bentwich, I., Avniel, A., Karov, Y., Aharonov, R., Gilad, S., Barad, O., Barzilai, A., Einat, P., Einav, U., Meiri, E., et al. (2005). Identification of hundreds of conserved and nonconserved human microRNAs. *Nat Genet* 37, 766–770.
- Blau, H.M., Chiu, C.P., and Webster, C. (1983). Cytoplasmic activation of human nuclear genes in stable heterocaryons. *Cell* 32, 1171–1180.

Blewitt, M.E., Gendrel, A.-V., Pang, Z., Sparrow, D.B., Whitelaw, N., Craig, J.M., Apedaile, A., Hilton, D.J., Dunwoodie, S.L., Brockdorff, N., et al. (2008). SmcHD1, containing a structural-maintenance-of-chromosomes hinge domain, has a critical role in X inactivation. *Nat. Genet.* 40, 663–669.

Block, G.J., Petek, L.M., Narayanan, D., Amell, A.M., Moore, J.M., Rabaia, N.A., Tyler, A., van der Maarel, S.M., Tawil, R., Filippova, G.N., et al. (2012). Asymmetric bidirectional transcription from the FSHD-causing D4Z4 array modulates DUX4 production. *PLoS One* 7, e35532.

Block, G.J., Narayanan, D., Amell, A.M., Petek, L.M., Davidson, K.C., Bird, T.D., Tawil, R., Moon, R.T., and Miller, D.G. (2013). Wnt/beta-catenin signaling suppresses DUX4 expression and prevents apoptosis of FSHD muscle cells. *Hum Mol Genet* 22, 4661–4672.

Bodega, B., Ramirez, G.D., Grasser, F., Cheli, S., Brunelli, S., Mora, M., Meneveri, R., Marozzi, A., Mueller, S., Battaglioli, E., et al. (2009). Remodeling of the chromatin structure of the facioscapulohumeral muscular dystrophy (FSHD) locus and upregulation of FSHD-related gene 1 (FRG1) expression during human myogenic differentiation. *BMC Biol* 7, 41.

Bohnsack, M.T., Czaplinski, K., and Gorlich, D. (2004). Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *Rna* 10, 185–191.

Bosnakovski, D., Xu, Z., Gang, E.J., Galindo, C.L., Liu, M., Simsek, T., Garner, H.R., Agha-Mohammadi, S., Tassin, A., Coppee, F., et al. (2008a). An isogenetic myoblast expression screen identifies DUX4-mediated FSHD-associated molecular pathologies. *Embo J* 27, 2766–2779.

Bosnakovski, D., Lamb, S., Simsek, T., Xu, Z., Belayew, A., Perlingeiro, R., and Kyba, M. (2008b). DUX4c, an FSHD candidate gene, interferes with myogenic regulators and abolishes myoblast differentiation. *Exp Neurol* 214, 87–96.

Braun, T., and Gautel, M. (2011). Transcriptional mechanisms regulating skeletal muscle differentiation, growth and homeostasis. *Nat. Rev. Mol. Cell Biol.* 12, 349–361.

Braun, G.S., Kretzler, M., Heider, T., Floege, J., Holzman, L.B., Kriz, W., and Moeller, M.J. (2007). Differentially spliced isoforms of FAT1 are asymmetrically distributed within migrating cells. *J. Biol. Chem.* 282, 22823–22833.

Broucqsault, N., Morere, J., Gaillard, M.C., Dumonceaux, J., Torrents, J., Salort-Campana, E., Maues De Paula, A., Bartoli, M., Fernandez, C., Chesnais, A.L., et al. (2013). Dysregulation of 4q35- and muscle-specific genes in fetuses with a short D4Z4 array linked to facio-scapulo-humeral dystrophy. *Hum Mol Genet* 22, 4206–4214.

Brouwer, O.F., Padberg, G.W., Ruys, C.J., Brand, R., de Laat, J.A., and Grote, J.J. (1991). Hearing loss in facioscapulohumeral muscular dystrophy. *Neurology* 41, 1878–1881.

Brouwer, O.F., Padberg, G.W., van der Ploeg, R.J., Ruys, C.J., and Brand, R. (1992). The influence of handedness on the distribution of muscular weakness of the arm in facioscapulohumeral muscular dystrophy. *Brain* 115 (Pt 5), 1587–1598.

Buckingham, M., and Rigby, P.W.J. (2014). Gene Regulatory Networks and Transcriptional Mechanisms that Control Myogenesis. *Dev. Cell* 28, 225–238.

Buongiorno-Nardelli, M., Micheli, G., Carri, M.T., and Marilley, M. (1982). A relationship between replicon size and supercoiled loop domains in the eukaryotic genome. *Nature* 298, 100–102.

Butler-Browne, G.S., Barbet, J.P., and Thornell, L.E. (1990). Myosin heavy and light chain expression during human skeletal muscle development and precocious muscle maturation induced by thyroid hormone. *Anat. Embryol. (Berl.)* 181, 513–522.

Cabianca, D.S., Casa, V., Bodega, B., Xynos, A., Ginelli, E., Tanaka, Y., and Gabellini, D. (2012). A long ncRNA links copy number variation to a polycomb/trithorax epigenetic switch in FSHD muscular dystrophy. *Cell* 149, 819–831.

Cacurri, S., Piazzo, N., Deidda, G., Vigneti, E., Galluzzi, G., Colantoni, L., Merico, B., Ricci, E., and Felicetti, L. (1998). Sequence Homology between 4qter and 10qter Loci Facilitates the Instability of Subtelomeric KpnI Repeat Units Implicated in Facioscapulohumeral Muscular Dystrophy. *Am. J. Hum. Genet.* 63, 181–190.

Caruso, N., Herberth, B., Bartoli, M., Puppo, F., Dumonceaux, J., Zimmermann, A., Denadai, S., Lebosse, M., Roche, S., Geng, L., et al. (2013). Deregulation of the protocadherin gene FAT1 alters muscle shapes: implications for the pathogenesis of facioscapulohumeral dystrophy. *PLoS Genet* 9, e1003550.

Celegato, B., Capitanio, D., Pescatori, M., Romualdi, C., Pacchioni, B., Cagnin, S., Vigano, A., Colantoni, L., Begum, S., Ricci, E., et al. (2006). Parallel protein and transcript profiles of FSHD patient muscles correlate to the D4Z4 arrangement and reveal a common impairment of slow to fast fibre différenciation and a general deregulation of MyoD-dependent genes. *Proteomics* 6, 5303–5321.

Chadwick, B.P. (2008). DXZ4 chromatin adopts an opposing conformation to that of the surrounding chromosome and acquires a novel inactive X-specific role involving CTCF and antisense transcripts. *Genome Res.* 18, 1259–1269.

Cho, D.H., Thienes, C.P., Mahoney, S.E., Analau, E., Filippova, G.N., and Tapscott, S.J. (2005). Antisense transcription and heterochromatin at the DM1 CTG repeats are constrained by CTCF. *Mol Cell* 20, 483–489.

Ciani, L., Patel, A., Allen, N.D., and ffrench-Constant, C. (2003). Mice lacking the giant protocadherin mFAT1 exhibit renal slit junction abnormalities and a partially penetrant cyclopia and anophthalmia phenotype. *Mol. Cell. Biol.* 23, 3575–3582.

Clapp, J., Mitchell, L.M., Bolland, D.J., Fantes, J., Corcoran, A.E., Scotting, P.J., Armour, J.A., and Hewitt, J.E. (2007). Evolutionary conservation of a coding function for D4Z4, the tandem DNA repeat mutated in facioscapulohumeral muscular dystrophy. *Am J Hum Genet* 81, 264–279.

Cockerill, P.N., and Garrard, W.T. (1986). Chromosomal loop anchorage of the kappa immunoglobulin gene occurs next to the enhancer in a region containing topoisomerase II sites. *Cell* 44, 273–282.

Collins, C.A., and Partridge, T.A. (2005). Self-renewal of the adult skeletal muscle satellite cell. *Cell Cycle Georget. Tex* 4, 1338–1341.

Corona, E.D., Jacquelin, D., Gatica, L., and Rosa, A.L. (2013). Multiple protein domains contribute to nuclear import and cell toxicity of DUX4, a candidate pathogenic protein for facioscapulohumeral muscular dystrophy. *PLoS One* 8, e75614.

Cremer, T., and Cremer, C. (2001). Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat. Rev. Genet.* 2, 292–301.

Damianov, A., and Black, D.L. (2010). Autoregulation of Fox protein expression to produce dominant negative splicing factors. *RNA N. Y. N* 16, 405–416.

Deak, K.L., Lemmers, R.J., Stajich, J.M., Klooster, R., Tawil, R., Frants, R.R., Speer, M.C., van der Maarel, S.M., and Gilbert, J.R. (2007). Genotype-phenotype study in an FSHD family with a proximal deletion encompassing p13E-11 and D4Z4. *Neurology* 68, 578–582.

Deidda, G., Cacurri, S., Grisanti, P., Vigneti, E., Piazzo, N., and Felicetti, L. (1995). Physical mapping evidence for a duplicated region on chromosome 10qter showing high homology with the facioscapulohumeral muscular dystrophy locus on chromosome 4qter. *Eur J Hum Genet* 3, 155–167.

Denli, A.M., Tops, B.B., Plasterk, R.H., Ketting, R.F., and Hannon, G.J. (2004). Processing of primary microRNAs by the Microprocessor complex. *Nature* 432, 231–235.

Van Deutekom, J.C., Wijmenga, C., van Tienhoven, E.A., Gruter, A.M., Hewitt, J.E., Padberg, G.W., van Ommen, G.J., Hofker, M.H., and Frants, R.R. (1993). FSHD associated DNA rearrangements are due to deletions of integral copies of a 3.2 kb tandemly repeated unit. *Hum Mol Genet* 2, 2037–2042.

Van Deutekom, J.C., Bakker, E., Lemmers, R.J., van der Wielen, M.J., Bik, E., Hofker, M.H., Padberg, G.W., and Frants, R.R. (1996). Evidence for subtelomeric exchange of 3.3 kb tandemly repeated units between chromosomes 4q35 and 10q26: implications for genetic counselling and etiology of FSHD1. *Hum Mol Genet* 5, 1997–2003.

Dixit, M., Anseau, E., Tassin, A., Winokur, S., Shi, R., Qian, H., Sauvage, S., Matteotti, C., van Acker, A.M., Leo, O., et al. (2007). DUX4, a candidate gene of facioscapulohumeral muscular dystrophy, encodes a transcriptional activator of PITX1. *Proc Natl Acad Sci U A* 104, 18157–18162.

Dmitriev, P., Petrov, A., Anseau, E., Stankevics, L., Charron, S., Kim, E., Bos, T.J., Robert, T., Turki, A., Coppee, F., et al. (2011). The Kruppel-like factor 15 as a molecular link between myogenic factors and a chromosome 4q transcriptional enhancer implicated in facioscapulohumeral dystrophy. *J Biol Chem* 286, 44620–44631.

Dunne, J., Hanby, A.M., Poulson, R., Jones, T.A., Sheer, D., Chin, W.G., Da, S.M., Zhao, Q., Beverley, P.C., and Owen, M.J. (1995). Molecular cloning and tissue expression of FAT, the human homologue of the *Drosophila* fat gene that is located on chromosome 4q34-q35 and encodes a putative adhesion molecule. *Genomics* 30, 207–223.

- Ehrlich, M., and Lacey, M. (2012). Deciphering transcription dysregulation in FSH muscular dystrophy. *J Hum Genet* 57, 477–484.
- Faustmann, P.M., Farahati, J., Rupilius, B., Dux, R., Koch, M.C., and Reiners, C. (1996). Cardiac involvement in facio-scapulo-humeral muscular dystrophy: a family study using Thallium-201 single-photon-emission-computed tomography. *J Neurol Sci* 144, 59–63.
- Ferreboeuf, M., Mariot, V., Bessieres, B., Vasiljevic, A., Attie-Bitach, T., Collardeau, S., Morere, J., Roche, S., Magdinier, F., Robin-Ducellier, J., et al. (2013). DUX4 and DUX4 downstream target genes are expressed in fetal FSHD muscles. *Hum Mol Genet* 23, 171–181.
- Flynt, A.S., Greimann, J.C., Chung, W.-J., Lima, C.D., and Lai, E.C. (2010). MicroRNA Biogenesis via Splicing and Exosome-Mediated Trimming in *Drosophila*. *Mol. Cell* 38, 900–907.
- Francastel, C., Schübeler, D., Martin, D.I., and Groudine, M. (2000). Nuclear compartmentalization and gene activity. *Nat. Rev. Mol. Cell Biol.* 1, 137–143.
- Friedman, R.C., Farh, K.K., Burge, C.B., and Bartel, D.P. (2009). Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res* 19, 92–105.
- Fukumura, K., Kato, A., Jin, Y., Ideue, T., Hirose, T., Kataoka, N., Fujiwara, T., Sakamoto, H., and Inoue, K. (2007). Tissue-specific splicing regulator Fox-1 induces exon skipping by interfering E complex formation on the downstream intron of human F1 gamma gene. *Nucleic Acids Res.* 35, 5303–5311.
- Gabellini, D., Green, M.R., and Tupler, R. (2002). Inappropriate gene activation in FSHD: a repressor complex binds a chromosomal repeat deleted in dystrophic muscle. *Cell* 110, 339–348.
- Gabellini, D., D’Antona, G., Moggio, M., PELLE, A., Zecca, C., Adami, R., Angeletti, B., Ciscato, P., Pellegrino, M.A., Bottinelli, R., et al. (2006). Facioscapulohumeral muscular dystrophy in mice overexpressing FRG1. *Nature* 439, 973–977.
- Galluzzi, G., Deidda, G., Cacurri, S., Colantoni, L., Piazzo, N., Vigneti, E., Ricci, E., Servidei, S., Merico, B., Pachi, A., et al. (1999). Molecular analysis of 4q35 rearrangements in facioscapulohumeral muscular dystrophy (FSHD): application to family studies for a correct genetic advice and a reliable prenatal diagnosis of the disease. *Neuromuscul Disord* 9, 190–198.
- Van Geel, M., Heather, L.J., Lyle, R., Hewitt, J.E., Frants, R.R., and de Jong, P.J. (1999). The FSHD region on human chromosome 4q35 contains potential coding regions among pseudogenes and a high density of repeat elements. *Genomics* 61, 55–65.
- Van Geel, M., Dickson, M.C., Beck, A.F., Bolland, D.J., Frants, R.R., van der Maarel, S.M., de Jong, P.J., and Hewitt, J.E. (2002). Genomic analysis of human chromosome 10q and 4q telomeres suggests a common origin. *Genomics* 79, 210–217.
- Gendrel, A.-V., Apedaile, A., Coker, H., Termanis, A., Zvetkova, I., Godwin, J., Tang, Y.A., Huntley, D., Montana, G., Taylor, S., et al. (2012). Smchd1-dependent and -independent

pathways determine developmental dynamics of CpG island methylation on the inactive X chromosome. *Dev. Cell* 23, 265–279.

Geng, L.N., Tyler, A.E., and Tapscott, S.J. (2011). Immunodetection of human double homeobox 4. *Hybrid. Larchmt* 30, 125–130.

Geng, L.N., Yao, Z., Snider, L., Fong, A.P., Cech, J.N., Young, J.M., van der Maarel, S.M., Ruzzo, W.L., Gentleman, R.C., Tawil, R., et al. (2012). DUX4 activates germline genes, retroelements, and immune mediators: implications for facioscapulohumeral dystrophy. *Dev Cell* 22, 38–51.

Gordon, S., Akopyan, G., Garban, H., and Bonavida, B. (2006). Transcription factor YY1: structure, function, and therapeutic implications in cancer biology. *Oncogene* 25, 1125–1142.

De Greef, J.C., Lemmers, R.J., van Engelen, B.G., Sacconi, S., Venance, S.L., Frants, R.R., Tawil, R., and van der Maarel, S.M. (2009). Common epigenetic changes of D4Z4 in contraction-dependent and contraction-independent FSHD. *Hum Mutat* 30, 1449–1459.

De Greef, J.C., Lemmers, R.J., Camano, P., Day, J.W., Sacconi, S., Dunand, M., van Engelen, B.G., Kiuru-Enari, S., Padberg, G.W., Rosa, A.L., et al. (2010). Clinical features of facioscapulohumeral muscular dystrophy 2. *Neurology* 75, 1548–1554.

Gregory, R.I., Yan, K.P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., and Shiekhattar, R. (2004). The Microprocessor complex mediates the genesis of microRNAs. *Nature* 432, 235–240.

Gwizdek, C., Ossareh-Nazari, B., Brownawell, A.M., Evers, S., Macara, I.G., and Dargemont, C. (2004). Minihelix-containing RNAs mediate exportin-5-dependent nuclear export of the double-stranded RNA-binding protein ILF3. *J Biol Chem* 279, 884–891.

Han, J., Lee, Y., Yeom, K.H., Kim, Y.K., Jin, H., and Kim, V.N. (2004). The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev* 18, 3016–3027.

Han, J., Lee, Y., Yeom, K.H., Nam, J.W., Heo, I., Rhee, J.K., Sohn, S.Y., Cho, Y., Zhang, B.T., and Kim, V.N. (2006). Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell* 125, 887–901.

Hanel, M.L., Wuebbles, R.D., and Jones, P.L. (2009). Muscular dystrophy candidate gene FRG1 is critical for muscle development. *Dev Dyn* 238, 1502–1512.

Harold, C. (2006). Atlas of Genetic diagnosis and counseling. Hum. Press.

Hewitt, J.E., Lyle, R., Clark, L.N., Valleley, E.M., Wright, T.J., Wijmenga, C., van Deutekom, J.C., Francis, F., Sharpe, P.T., Hofker, M., et al. (1994). Analysis of the tandem repeat locus D4Z4 associated with facioscapulohumeral muscular dystrophy. *Hum Mol Genet* 3, 1287–1295.

Jackson, D.A. (2003). The principles of nuclear structure. *Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol.* 11, 387–401.

Jacqueline, M. (2011). The identification of Landouzy-Dejerine disease: an investigative history. *Neuromuscul Disord* 21, 291–297.

- Jiang, G., Yang, F., van Overveld, P.G., Vedanarayanan, V., van der Maarel, S., and Ehrlich, M. (2003). Testing the position-effect variegation hypothesis for facioscapulohumeral muscular dystrophy by analysis of histone modification and gene expression in subtelomeric 4q. *Hum Mol Genet* 12, 2909–2921.
- Jones, T.I., Chen, J.C., Rahimov, F., Homma, S., Arashiro, P., Beermann, M.L., King, O.D., Miller, J.B., Kunkel, L.M., Emerson, C.P., Jr., et al. (2012). Facioscapulohumeral muscular dystrophy family studies of DUX4 expression: evidence for disease modifiers and a quantitative model of pathogenesis. *Hum Mol Genet* 21, 4419–4430.
- Kalkman, J.S., Schillings, M.L., Zwarts, M.J., van Engelen, B.G., and Bleijenberg, G. (2006). Influence of relatives on fatigue experienced by patients with facioscapulohumeral dystrophy, myotonic dystrophy and HMSN-I. *Eur Neurol* 56, 24–30.
- Kalkman, J.S., Zwarts, M.J., Schillings, M.L., van Engelen, B.G., and Bleijenberg, G. (2008). Different types of fatigue in patients with facioscapulohumeral dystrophy, myotonic dystrophy and HMSN-I. Experienced fatigue and physiological fatigue. *Neurol Sci* 29 Suppl 2, S238–40.
- Katoh, Y., and Katoh, M. (2006). Comparative integromics on FAT1, FAT2, FAT3 and FAT4. *Int. J. Mol. Med.* 18, 523–528.
- Khvorova, A., Reynolds, A., and Jayasena, S.D. (2003). Functional siRNAs and miRNAs exhibit strand bias. *Cell* 115, 209–216.
- Kim, Y.K., and Kim, V.N. (2007). Processing of intronic microRNAs. *Embo J* 26, 775–783.
- Kim, S.K., Lund, J., Kiraly, M., Duke, K., Jiang, M., Stuart, J.M., Eizinger, A., Wylie, B.N., and Davidson, G.S. (2001). A gene expression map for *Caenorhabditis elegans*. *Science* 293, 2087–2092.
- Van Koningsbruggen, S., Dirks, R.W., Mommaas, A.M., Onderwater, J.J., Deidda, G., Padberg, G.W., Frants, R.R., and van der Maarel, S.M. (2004). FRG1P is localised in the nucleolus, Cajal bodies, and speckles. *J. Med. Genet.* 41, e46.
- Kowaljow, V., Marcowycz, A., Ansseau, E., Conde, C.B., Sauvage, S., Matteotti, C., Arias, C., Corona, E.D., Nunez, N.G., Leo, O., et al. (2007). The DUX4 gene at the FSHD1A locus encodes a pro-apoptotic protein. *Neuromuscul Disord* 17, 611–623.
- Krol, J., Sobczak, K., Wilczynska, U., Drath, M., Jasinska, A., Kaczynska, D., and Krzyzosiak, W.J. (2004). Structural features of microRNA (miRNA) precursors and their relevance to miRNA biogenesis and small interfering RNA/short hairpin RNA design. *J Biol Chem* 279, 42230–42239.
- Krom, Y.D., Dumonceaux, J., Mamchaoui, K., den Hamer, B., Mariot, V., Negroni, E., Geng, L.N., Martin, N., Tawil, R., Tapscott, S.J., et al. (2012). Generation of isogenic D4Z4 contracted and noncontracted immortal muscle cell clones from a mosaic patient: a cellular model for FSHD. *Am J Pathol* 181, 1387–1401.

- Krom, Y.D., Thijssen, P.E., Young, J.M., den Hamer, B., Balog, J., Yao, Z., Maves, L., Snider, L., Knopp, P., Zammit, P.S., et al. (2013). Intrinsic epigenetic regulation of the D4Z4 macrosatellite repeat in a transgenic mouse model for FSHD. *PLoS Genet* 9, e1003415.
- Kuang, S., Kuroda, K., Le Grand, F., and Rudnicki, M.A. (2007). Asymmetric self-renewal and commitment of satellite stem cells in muscle. *Cell* 129, 999–1010.
- Kuang, S., Gillespie, M.A., and Rudnicki, M.A. (2008). Niche regulation of muscle satellite cell self-renewal and differentiation. *Cell Stem Cell* 2, 22–31.
- Kuroyanagi, H. (2009). Fox-1 family of RNA-binding proteins. *Cell. Mol. Life Sci. CMLS* 66, 3895–3907.
- Lachner, M., O’Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* 410, 116–120.
- Lagha, M., Sato, T., Bajard, L., Daubas, P., Esner, M., Montarras, D., Relaix, F., and Buckingham, M. (2008). Regulation of skeletal muscle stem cell behavior by Pax3 and Pax7. *Cold Spring Harb. Symp. Quant. Biol.* 73, 307–315.
- Lagos-Quintana, M., Rauhut, R., Lendeckel, W., and Tuschl, T. (2001). Identification of novel genes coding for small expressed RNAs. *Science* 294, 853–858.
- Landthaler, M., Yalcin, A., and Tuschl, T. (2004). The human DiGeorge syndrome critical region gene 8 and Its D. melanogaster homolog are required for miRNA biogenesis. *Curr Biol* 14, 2162–2167.
- Laoudj-Chenivesse, D., Carnac, G., Bisbal, C., Hugon, G., Bouillot, S., Desnuelle, C., Vassetzky, Y., and Fernandez, A. (2005). Increased levels of adenine nucleotide translocator 1 protein and response to oxidative stress are early events in facioscapulohumeral muscular dystrophy muscle. *J Mol Med Berl* 83, 216–224.
- Lassche, S., Stienen, G.J., Irving, T.C., van der Maarel, S.M., Voermans, N.C., Padberg, G.W., Granzier, H., van Engelen, B.G., and Ottenheijm, C.A. (2013). Sarcomeric dysfunction contributes to muscle weakness in facioscapulohumeral muscular dystrophy. *Neurology* 80, 733–737.
- Lau, N.C., Lim, L.P., Weinstein, E.G., and Bartel, D.P. (2001). An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* 294, 858–862.
- Lee, R.C., and Ambros, V. (2001). An extensive class of small RNAs in *Caenorhabditis elegans*. *Science* 294, 862–864.
- Lee, J.H., Goto, K., Matsuda, C., and Arahata, K. (1995). Characterization of a tandemly repeated 3.3-kb KpnI unit in the facioscapulohumeral muscular dystrophy (FSHD) gene region on chromosome 4q35. *Muscle Nerve Suppl* 2, S6–13.
- Lee, R.C., Feinbaum, R.L., and Ambros, V. (1993). The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* 75, 843–854.
- Lee, Y., Jeon, K., Lee, J.T., Kim, S., and Kim, V.N. (2002). MicroRNA maturation: stepwise processing and subcellular localization. *Embo J* 21, 4663–4670.

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Radmark, O., Kim, S., et al. (2003). The nuclear RNase III Drosha initiates microRNA processing. *Nature* 425, 415–419.

Lee, Y., Kim, M., Han, J., Yeom, K.H., Lee, S., Baek, S.H., and Kim, V.N. (2004). MicroRNA genes are transcribed by RNA polymerase II. *Embo J* 23, 4051–4060.

Lemmers, R.J., van der Maarel, S.M., van Deutekom, J.C., van der Wielen, M.J., Deidda, G., Dauwerse, H.G., Hewitt, J., Hofker, M., Bakker, E., Padberg, G.W., et al. (1998). Inter- and intrachromosomal sub-telomeric rearrangements on 4q35: implications for facioscapulohumeral muscular dystrophy (FSHD) aetiology and diagnosis. *Hum Mol Genet* 7, 1207–1214.

Lemmers, R.J., de Kievit, P., Sandkuijl, L., Padberg, G.W., van Ommen, G.J., Frants, R.R., and van der Maarel, S.M. (2002). Facioscapulohumeral muscular dystrophy is uniquely associated with one of the two variants of the 4q subtelomere. *Nat Genet* 32, 235–236.

Lemmers, R.J., Osborn, M., Haaf, T., Rogers, M., Frants, R.R., Padberg, G.W., Cooper, D.N., van der Maarel, S.M., and Upadhyaya, M. (2003). D4F104S1 deletion in facioscapulohumeral muscular dystrophy: phenotype, size, and detection. *Neurology* 61, 178–183.

Lemmers, R.J., van der Wielen, M.J., Bakker, E., Padberg, G.W., Frants, R.R., and van der Maarel, S.M. (2004). Somatic mosaicism in FSHD often goes undetected. *Ann Neurol* 55, 845–850.

Lemmers, R.J., Wohlgemuth, M., van der Gaag, K.J., van der Vliet, P.J., van Teijlingen, C.M., de Knijff, P., Padberg, G.W., Frants, R.R., and van der Maarel, S.M. (2007). Specific sequence variations within the 4q35 region are associated with facioscapulohumeral muscular dystrophy. *Am J Hum Genet* 81, 884–894.

Lemmers, R.J., van der Vliet, P.J., Klooster, R., Sacconi, S., Camano, P., Dauwerse, J.G., Snider, L., Straasheijm, K.R., van Ommen, G.J., Padberg, G.W., et al. (2010). A unifying genetic model for facioscapulohumeral muscular dystrophy. *Science* 329, 1650–1653.

Lemmers, R.J., Tawil, R., Petek, L.M., Balog, J., Block, G.J., Santen, G.W., Amell, A.M., van der Vliet, P.J., Almomani, R., Straasheijm, K.R., et al. (2012). Digenic inheritance of an SMCHD1 mutation and an FSHD-permissive D4Z4 allele causes facioscapulohumeral muscular dystrophy type 2. *Nat Genet* 44, 1370–1374.

Lim, L.P., Lau, N.C., Weinstein, E.G., Abdelhakim, A., Yekta, S., Rhoades, M.W., Burge, C.B., and Bartel, D.P. (2003). The microRNAs of *Caenorhabditis elegans*. *Genes Dev* 17, 991–1008.

Lim, L.P., Lau, N.C., Garrett-Engle, P., Grimson, A., Schelter, J.M., Castle, J., Bartel, D.P., Linsley, P.S., and Johnson, J.M. (2005). Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* 433, 769–773.

Liu, Q., Jones, T.I., Tang, V.W., Briher, W.M., and Jones, P.L. (2010). Facioscapulohumeral muscular dystrophy region gene-1 (FRG-1) is an actin-bundling protein associated with muscle-attachment sites. *J Cell Sci* 123, 1116–1123.

- Luco, R.F., Pan, Q., Tominaga, K., Blencowe, B.J., Pereira-Smith, O.M., and Misteli, T. (2010). Regulation of alternative splicing by histone modifications. *Science* 327, 996–1000.
- Lunt, P.W., and Harper, P.S. (1991). Genetic counselling in facioscapulohumeral muscular dystrophy. *J Med Genet* 28, 655–664.
- Lunt, P.W., Compston, D.A., and Harper, P.S. (1989). Estimation of age dependent penetrance in facioscapulohumeral muscular dystrophy by minimising ascertainment bias. *J Med Genet* 26, 755–760.
- Lunt, P.W., Jardine, P.E., Koch, M.C., Maynard, J., Osborn, M., Williams, M., Harper, P.S., and Upadhyaya, M. (1995). Correlation between fragment size at D4F104S1 and age at onset or at wheelchair use, with a possible generational effect, accounts for much phenotypic variation in 4q35-facioscapulohumeral muscular dystrophy (FSHD). *Hum Mol Genet* 4, 951–958.
- Van der Maarel, S.M., Deidda, G., Lemmers, R.J., van Overveld, P.G., van der Wielen, M., Hewitt, J.E., Sandkuijl, L., Bakker, B., van Ommen, G.J., Padberg, G.W., et al. (2000). De novo facioscapulohumeral muscular dystrophy: frequent somatic mosaicism, sex-dependent phenotype, and the role of mitotic transchromosomal repeat interaction between chromosomes 4 and 10. *Am J Hum Genet* 66, 26–35.
- Van der Maarel, S.M., Frants, R.R., and Padberg, G.W. (2007). Facioscapulohumeral muscular dystrophy. *Biochim Biophys Acta* 1772, 186–194.
- Mamchaoui, K., Trollet, C., Bigot, A., Negroni, E., Chaouch, S., Wolff, A., Kandalla, P.K., Marie, S., Di Santo, J., St Guily, J.L., et al. (2011). Immortalized pathological human myoblasts: towards a universal tool for the study of neuromuscular disorders. *Skelet. Muscle* 1, 34.
- Marshall, W.F. (2003). Gene expression and nuclear architecture during development and differentiation. *Mech. Dev.* 120, 1217–1230.
- Masny, P.S., Chan, O.Y., de Greef, J.C., Bengtsson, U., Ehrlich, M., Tawil, R., Lock, L.F., Hewitt, J.E., Stocksdales, J., Martin, J.H., et al. (2010). Analysis of allele-specific RNA transcription in FSHD by RNA-DNA FISH in single myonuclei. *Eur J Hum Genet* 18, 448–456.
- McKee, A.E., Minet, E., Stern, C., Riahi, S., Stiles, C.D., and Silver, P.A. (2005). A genome-wide in situ hybridization map of RNA-binding proteins reveals anatomically restricted expression in the developing mouse brain. *BMC Dev. Biol.* 5, 14.
- Migeon, B.R. (2006). The role of X inactivation and cellular mosaicism in women's health and sex-specific diseases. *JAMA J. Am. Med. Assoc.* 295, 1428–1433.
- Mirkovitch, J., Mirault, M.E., and Laemmli, U.K. (1984). Organization of the higher-order chromatin loop: specific DNA attachment sites on nuclear scaffold. *Cell* 39, 223–232.
- Mitsuhashi, H., Mitsuhashi, S., Lynn-Jones, T., Kawahara, G., and Kunkel, L.M. (2012). Expression of DUX4 in zebrafish development recapitulates facioscapulohumeral muscular dystrophy. *Hum Mol Genet* 22, 568–577.

- Morlando, M., Ballarino, M., Gromak, N., Pagano, F., Bozzoni, I., and Proudfoot, N.J. (2008). Primary microRNA transcripts are processed co-transcriptionally. *Nat. Struct. Mol. Biol.* *15*, 902–909.
- Mourelatos, Z., Dostie, J., Paushkin, S., Sharma, A., Charroux, B., Abel, L., Rappsilber, J., Mann, M., and Dreyfuss, G. (2002). miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs. *Genes Dev* *16*, 720–728.
- Nguyen, K., Walrafen, P., Bernard, R., Attarian, S., Chaix, C., Vovan, C., Renard, E., Dufrane, N., Pouget, J., Vannier, A., et al. (2011). Molecular combing reveals allelic combinations in facioscapulohumeral dystrophy. *Ann Neurol* *70*, 627–633.
- Okamura, K., Phillips, M.D., Tyler, D.M., Duan, H., Chou, Y.T., and Lai, E.C. (2008). The regulatory activity of microRNA* species has substantial influence on microRNA and 3' UTR evolution. *Nat Struct Mol Biol* *15*, 354–363.
- Osborne, R.J., Welle, S., Venance, S.L., Thornton, C.A., and Tawil, R. (2007). Expression profile of FSHD supports a link between retinal vasculopathy and muscular dystrophy. *Neurology* *68*, 569–577.
- Ottaviani, A., Rival-Gervier, S., Boussouar, A., Foerster, A.M., Rondier, D., Sacconi, S., Desnuelle, C., Gilson, E., and Magdinier, F. (2009). The D4Z4 macrosatellite repeat acts as a CTCF and A-type lamins-dependent insulator in facio-scapulo-humeral dystrophy. *PLoS Genet* *5*, e1000394.
- Ottaviani, A., Schluth-Bolard, C., Gilson, E., and Magdinier, F. (2010). D4Z4 as a prototype of CTCF and lamins-dependent insulator in human cells. *Nucleus* *1*, 30–36.
- Van Overveld, P.G., Lemmers, R.J., Deidda, G., Sandkuijl, L., Padberg, G.W., Frants, R.R., and van der Maarel, S.M. (2000). Interchromosomal repeat array interactions between chromosomes 4 and 10: a model for subtelomeric plasticity. *Hum Mol Genet* *9*, 2879–2884.
- Van Overveld, P.G.M., Lemmers, R.J.F.L., Sandkuijl, L.A., Enthoven, L., Winokur, S.T., Bakels, F., Padberg, G.W., van Ommen, G.-J.B., Frants, R.R., and van der Maarel, S.M. (2003). Hypomethylation of D4Z4 in 4q-linked and non-4q-linked facioscapulohumeral muscular dystrophy. *Nat. Genet.* *35*, 315–317.
- Padberg, G.W., Brouwer, O.F., de Keizer, R.J., Dijkman, G., Wijmenga, C., Grote, J.J., and Frants, R.R. (1995). On the significance of retinal vascular disease and hearing loss in facioscapulohumeral muscular dystrophy. *Muscle Nerve Suppl* S73–80.
- Perry, R.L., and Rudnick, M.A. (2000). Molecular mechanisms regulating myogenic determination and differentiation. *Front. Biosci. J. Virtual Libr.* *5*, D750–767.
- Petrov, A., Laoudj, D., and Vasetskiĭ, E. (2003). [Genetics and epigenetics of facio-scapulohumeral progressive (Landouzy-Dejerine) muscular dystrophy]. *Genetika* *39*, 202–206.
- Petrov, A., Pirozhkova, I., Carnac, G., Laoudj, D., Lipinski, M., and Vassetzky, Y.S. (2006). Chromatin loop domain organization within the 4q35 locus in facioscapulohumeral dystrophy patients versus normal human myoblasts. *Proc Natl Acad Sci U S A* *103*, 6982–6987.

- Petrov, A., Allinne, J., Pirozhkova, I., Laoudj, D., Lipinski, M., and Vassetzky, Y.S. (2008). A nuclear matrix attachment site in the 4q35 locus has an enhancer-blocking activity in vivo: implications for the facio-scapulo-humeral dystrophy. *Genome Res* 18, 39–45.
- Pipes, G.C.T., Creemers, E.E., and Olson, E.N. (2006). The myocardin family of transcriptional coactivators: versatile regulators of cell growth, migration, and myogenesis. *Genes Dev.* 20, 1545–1556.
- Pirozhkova, I., Petrov, A., Dmitriev, P., Laoudj, D., Lipinski, M., and Vassetzky, Y. (2008). A functional role for 4qA/B in the structural rearrangement of the 4q35 region and in the regulation of FRG1 and ANT1 in facioscapulohumeral dystrophy. *PLoS One* 3, e3389.
- Ralston, E., and Hall, Z.W. (1989). Transfer of a protein encoded by a single nucleus to nearby nuclei in multinucleated myotubes. *Science* 244, 1066–1069.
- Ranum, L.P., and Day, J.W. (2004). Myotonic dystrophy: RNA pathogenesis comes into focus. *Am J Hum Genet* 74, 793–804.
- Rappsilber, J., Ryder, U., Lamond, A.I., and Mann, M. (2002). Large-scale proteomic analysis of the human spliceosome. *Genome Res.* 12, 1231–1245.
- Reed, P., Porter, N.C., Strong, J., Pumplin, D.W., Corse, A.M., Luther, P.W., Flanigan, K.M., and Bloch, R.J. (2006). Sarcolemmal reorganization in facioscapulohumeral muscular dystrophy. *Ann Neurol* 59, 289–297.
- Richards, M., Coppee, F., Thomas, N., Belayew, A., and Upadhyaya, M. (2012). Facioscapulohumeral muscular dystrophy (FSHD): an enigma unravelled? *Hum Genet* 131, 325–340.
- Rijkers, T., Deidda, G., van Koningsbruggen, S., van Geel, M., Lemmers, R.J., van Deutekom, J.C., Figlewicz, D., Hewitt, J.E., Padberg, G.W., Frants, R.R., et al. (2004). FRG2, an FSHD candidate gene, is transcriptionally upregulated in differentiating primary myoblast cultures of FSHD patients. *J Med Genet* 41, 826–836.
- Rodriguez, A., Griffiths-Jones, S., Ashurst, J.L., and Bradley, A. (2004). Identification of mammalian microRNA host genes and transcription units. *Genome Res* 14, 1902–1910.
- Ruby, J.G., Jan, C.H., and Bartel, D.P. (2007). Intronic microRNA precursors that bypass Drosha processing. *Nature* 448, 83–86.
- Sakellariou, P., Kekou, K., Fryssira, H., Sofocleous, C., Manta, P., Panousopoulou, A., Gounaris, K., and Kanavakis, E. (2012). Mutation spectrum and phenotypic manifestation in FSHD Greek patients. *Neuromuscul Disord* 22, 339–349.
- Salort-Campana, E., Nguyen, K., Levy, N., Pouget, J., and Attarian, S. (2012). [Clinical and molecular diagnosis of facioscapulohumeral dystrophy type 1 (FSHD1) in 2012]. *Rev Neurol Paris* 169, 573–582.
- Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc, B., and Cavalli, G. (2007). Genome regulation by polycomb and trithorax proteins. *Cell* 128, 735–745.

- Schwarz, D.S., Hutvagner, G., Du, T., Xu, Z., Aronin, N., and Zamore, P.D. (2003). Asymmetry in the assembly of the RNAi enzyme complex. *Cell* 115, 199–208.
- Scionti, I., Greco, F., Ricci, G., Govi, M., Arashiro, P., Vercelli, L., Berardinelli, A., Angelini, C., Antonini, G., Cao, M., et al. (2012). Large-scale population analysis challenges the current criteria for the molecular diagnosis of facioscapulohumeral muscular dystrophy. *Am J Hum Genet* 90, 628–635.
- Scott, M.P. (1992). Vertebrate homeobox gene nomenclature. *Cell* 71, 551–553.
- Sharer, J.D. (2005). The adenine nucleotide translocase type 1 (ANT1): a new factor in mitochondrial disease. *IUBMB Life* 57, 607–614.
- Sharma, V., Harafuji, N., Belayew, A., and Chen, Y.W. (2013). DUX4 differentially regulates transcriptomes of human rhabdomyosarcoma and mouse C2C12 cells. *PLoS One* 8, e64691.
- Shopland, L.S., Johnson, C.V., Byron, M., McNeil, J., and Lawrence, J.B. (2003). Clustering of multiple specific genes and gene-rich R-bands around SC-35 domains: evidence for local euchromatic neighborhoods. *J. Cell Biol.* 162, 981–990.
- Snider, L., Asawachaicharn, A., Tyler, A.E., Geng, L.N., Petek, L.M., Maves, L., Miller, D.G., Lemmers, R.J., Winokur, S.T., Tawil, R., et al. (2009). RNA transcripts, miRNA-sized fragments and proteins produced from D4Z4 units: new candidates for the pathophysiology of facioscapulohumeral dystrophy. *Hum Mol Genet* 18, 2414–2430.
- Snider, L., Geng, L.N., Lemmers, R.J., Kyba, M., Ware, C.B., Nelson, A.M., Tawil, R., Filippova, G.N., van der Maarel, S.M., Tapscott, S.J., et al. (2010). Facioscapulohumeral dystrophy: incomplete suppression of a retrotransposed gene. *PLoS Genet* 6, e1001181.
- Spector, D.L. (2001). Nuclear domains. *J. Cell Sci.* 114, 2891–2893.
- Srivastava, M., Fleming, P.J., Pollard, H.B., and Burns, A.L. (1989). Cloning and sequencing of the human nucleolin cDNA. *FEBS Lett.* 250, 99–105.
- Stadler, G., Rahimov, F., King, O.D., Chen, J.C., Robin, J.D., Wagner, K.R., Shay, J.W., Emerson, C.P., Jr., and Wright, W.E. (2013). Telomere position effect regulates DUX4 in human facioscapulohumeral muscular dystrophy. *Nat Struct Mol Biol* 20, 671–678.
- Stubgen, J.P. (2007). Facioscapulohumeral muscular dystrophy. A quantitative electromyographic study. *Electromyogr Clin Neurophysiol* 47, 175–182.
- Tam, W. (2001). Identification and characterization of human BIC, a gene on chromosome 21 that encodes a noncoding RNA. *Gene* 274, 157–167.
- Tam, R., Smith, K.P., and Lawrence, J.B. (2004). The 4q subtelomere harboring the FSHD locus is specifically anchored with peripheral heterochromatin unlike most human telomeres. *J Cell Biol* 167, 269–279.
- Tang, Z.Z., Zheng, S., Nikolic, J., and Black, D.L. (2009). Developmental control of CaV1.2 L-type calcium channel splicing by Fox proteins. *Mol. Cell. Biol.* 29, 4757–4765.
- Tanoue, T. (2005). New insights into Fat cadherins. *J. Cell Sci.* 118, 2347–2353.

Tassin, A., Laoudj-Chenivesse, D., Vanderplanck, C., Barro, M., Charron, S., Anseau, E., Chen, Y.W., Mercier, J., Coppee, F., and Belayew, A. (2013). DUX4 expression in FSHD muscle cells: how could such a rare protein cause a myopathy? *J Cell Mol Med* 17, 76–89.

Tonini, M.M., Passos-Bueno, M.R., Cerqueira, A., Mاتيoli, S.R., Pavanello, R., and Zatz, M. (2004). Asymptomatic carriers and gender differences in facioscapulohumeral muscular dystrophy (FSHD). *Neuromuscul Disord* 14, 33–38.

Tsumagari, K., Chang, S.C., Lacey, M., Baribault, C., Chittur, S.V., Sowden, J., Tawil, R., Crawford, G.E., and Ehrlich, M. (2011). Gene expression during normal and FSHD myogenesis. *BMC Med Genomics* 4, 67.

Tupler, R., Berardinelli, A., Barbierato, L., Frants, R., Hewitt, J.E., Lanzi, G., Maraschio, P., and Tiepolo, L. (1996). Monosomy of distal 4q does not cause facioscapulohumeral muscular dystrophy. *J Med Genet* 33, 366–370.

Turki, A., Hayot, M., Carnac, G., Pillard, F., Passerieux, E., Bommart, S., Raynaud de Mauverger, E., Hugon, G., Pincemail, J., Pietri, S., et al. (2012). Functional muscle impairment in facioscapulohumeral muscular dystrophy is correlated with oxidative stress and mitochondrial dysfunction. *Free Radic Biol Med* 53, 1068–1079.

Tyler, F. (1950). STUDIES IN DISORDERS OF MUSCLE. II. CLINICAL MANIFESTATIONS AND INHERITANCE OF FACIOSCAPULOHUMERAL DYSTROPHY IN A LARGE FAMILY. *Ann. Intern. Med.* 32, 640.

Upadhyaya, M., Lunt, P., Sarfarazi, M., Broadhead, W., Farnham, J., and Harper, P.S. (1992). The mapping of chromosome 4q markers in relation to facioscapulohumeral muscular dystrophy (FSHD). *Am J Hum Genet* 51, 404–410.

Wallace, L.M., Garwick, S.E., Mei, W., Belayew, A., Coppee, F., Ladner, K.J., Guttridge, D., Yang, J., and Harper, S.Q. (2011). DUX4, a candidate gene for facioscapulohumeral muscular dystrophy, causes p53-dependent myopathy in vivo. *Ann Neurol* 69, 540–552.

Wanschura, S., Schoenmakers, E.F., Huysmans, C., Bartnitzke, S., Van de Ven, W.J., and Bullrdiek, J. (1996). Mapping of the human HMG2 gene to 4q31. *Genomics* 31, 264–265.

Wightman, B., Ha, I., and Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in *C. elegans*. *Cell* 75, 855–862.

Wijmenga, C., Frants, R.R., Brouwer, O.F., Moerer, P., Weber, J.L., and Padberg, G.W. (1990). Location of facioscapulohumeral muscular dystrophy gene on chromosome 4. *Lancet* 336, 651–653.

Wijmenga, C., Padberg, G.W., Moerer, P., Wiegant, J., Liem, L., Brouwer, O.F., Milner, E.C., Weber, J.L., van Ommen, G.B., Sandkuijl, L.A., et al. (1991). Mapping of facioscapulohumeral muscular dystrophy gene to chromosome 4q35-qter by multipoint linkage analysis and in situ hybridization. *Genomics* 9, 570–575.

Wijmenga, C., Sandkuijl, L.A., Moerer, P., van der Boorn, N., Bodrug, S.E., Ray, P.N., Brouwer, O.F., Murray, J.C., van Ommen, G.J., Padberg, G.W., et al. (1992). Genetic linkage

map of facioscapulohumeral muscular dystrophy and five polymorphic loci on chromosome 4q35-qter. *Am J Hum Genet* 51, 411–415.

Winokur, S.T., Bengtsson, U., Feddersen, J., Mathews, K.D., Weiffenbach, B., Bailey, H., Markovich, R.P., Murray, J.C., Wasmuth, J.J., Altherr, M.R., et al. (1994). The DNA rearrangement associated with facioscapulohumeral muscular dystrophy involves a heterochromatin-associated repetitive element: implications for a role of chromatin structure in the pathogenesis of the disease. *Chromosome Res* 2, 225–234.

Winokur, S.T., Barrett, K., Martin, J.H., Forrester, J.R., Simon, M., Tawil, R., Chung, S.A., Masny, P.S., and Figlewicz, D.A. (2003). Facioscapulohumeral muscular dystrophy (FSHD) myoblasts demonstrate increased susceptibility to oxidative stress. *Neuromuscul Disord* 13, 322–333.

Wohlgemuth, M., van der Kooi, E.L., van Kesteren, R.G., van der Maarel, S.M., and Padberg, G.W. (2004). Ventilatory support in facioscapulohumeral muscular dystrophy. *Neurology* 63, 176–178.

Wu, S.-L., Tsai, M.-S., Wong, S.-H., Hsieh-Li, H.-M., Tsai, T.-S., Chang, W.-T., Huang, S.-L., Chiu, C.-C., and Wang, S.-H. (2010). Characterization of genomic structures and expression profiles of three tandem repeats of a mouse double homeobox gene: *Duxbl*. *Dev. Dyn.* 239, 927–940.

Wuebbles, R.D., Hanel, M.L., and Jones, P.L. (2009). FSHD region gene 1 (FRG1) is crucial for angiogenesis linking FRG1 to facioscapulohumeral muscular dystrophy-associated vasculopathy. *Model Mech* 2, 267–274.

Wuebbles, R.D., Long, S.W., Hanel, M.L., and Jones, P.L. (2010). Testing the effects of FSHD candidate gene expression in vertebrate muscle development. *Int J Clin Exp Pathol* 3, 386–400.

Yi, R., Qin, Y., Macara, I.G., and Cullen, B.R. (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev* 17, 3011–3016.

Zatz, M., Marie, S.K., Passos-Bueno, M.R., Vainzof, M., Campiotto, S., Cerqueira, A., Wijmenga, C., Padberg, G., and Frants, R. (1995). High proportion of new mutations and possible anticipation in Brazilian facioscapulohumeral muscular dystrophy families. *Am J Hum Genet* 56, 99–105.

Zeng, Y., and Cullen, B.R. (2004). Structural requirements for pre-microRNA binding and nuclear export by Exportin 5. *Nucleic Acids Res* 32, 4776–4785.

Zeng, W., de Greef, J.C., Chen, Y.Y., Chien, R., Kong, X., Gregson, H.C., Winokur, S.T., Pyle, A., Robertson, K.D., Schmiesing, J.A., et al. (2009). Specific loss of histone H3 lysine 9 trimethylation and HP1gamma/cohesin binding at D4Z4 repeats is associated with facioscapulohumeral dystrophy (FSHD). *PLoS Genet* 5, e1000559.

ANNEXES

PUBLICATION N°3: Dysregulation of 4q35- and muscle-specific gene in fetuses with a short D4Z4 array linked to facio-scapulo-humeral dystrophy

Broucqsault N, Morere J, Gaillard MC, Dumonceaux J, Torrents J, Salort-Campana E, Maues De Paula A, Bartoli M, Fernandez C, Chesnais AL, **Ferreboeuf M**, Sarda L, Dufour H, Desnuelle C, Attarian S, Levy N, Nguyen K, Magdinier F, Roche S. *Dysregulation of 4q35- and muscle-specific genes in fetuses with a short D4Z4 array linked to facio-scapulo-humeral dystrophy*. Hum Mol Genet. 2013 Oct 15; 22(20): 4206-14. doi: 10.1093/hmg/ddt272. Epub 2013 Jun 17.

Résumé:

Cette étude à laquelle nous avons participé a utilisé une plus grande cohorte de biopsies fœtales FSHD et Contrôle afin de déterminer s'il existe des changements majeurs dans l'expression de gènes impliqués dans la myogenèse ou situés sur le locus FSHD en 4q35.

Pour cela, des biopsies musculaires provenant de 6 fœtus porteurs de contractions *D4Z4*, âgés de 16 à 26 semaines de développements, ont été comparées à 21 biopsies fœtales contrôles situées dans la même tranche d'âge (12 à 37 semaines). Dans un premier temps, un grand nombre de gènes, impliqués dans la myogenèse, ont été identifiés comme spécifiquement dérégulés chez les fœtus porteurs de contractions, suggérant qu'à des stades précoces, un certain nombre de défauts musculaires pourraient être initié et entraîner l'apparition de la pathologie. Les mêmes analyses ont été effectuées sur des biopsies adultes FSHD et Contrôles mais aucune différence significative n'a été identifiée.

Dans un second temps, l'analyse de l'expression de 36 gènes du locus 4q35 a révélé que 14 étaient dérégulés spécifiquement au stade fœtal. Enfin, l'expression de *DUX4-FL* a été confirmée dans des biopsies fœtales et adultes FSHD mais également à une moindre fréquence dans des biopsies contrôles.

L'ensemble de ces résultats démontre donc que chez des fœtus porteurs de contractions *D4Z4*, en dehors de toute atteinte musculaire, un grand nombre de gènes du locus 4q35 ou impliqués dans la myogénèse sont dérégulés. Ainsi, bien que la FSHD soit une pathologie se déclarant à l'adolescence, des défauts moléculaires sont déjà présents à des stades précoces pouvant participer à la physiopathologie de la FSHD

Dysregulation of 4q35- and muscle-specific genes in fetuses with a short D4Z4 array linked to facio-scapulo-humeral dystrophy

Natacha Broucqsault^{1,†}, Julia Morere^{1,†}, Marie-Cécile Gaillard¹, Julie Dumonceaux⁶, Julia Torrents², Emmanuelle Salort-Campana^{1,3}, André Maues De Paula^{1,2}, Marc Bartoli¹, Carla Fernandez², Anne Laure Chesnais², Maxime Ferreboeuf⁶, Laure Sarda², Henry Dufour³, Claude Desnuelle⁷, Shahram Attarian^{1,4}, Nicolas Levy^{1,5}, Karine Nguyen^{1,5}, Frédéric Magdinier^{1,†,*} and Stéphane Roche^{1,†}

¹Aix Marseille Université, INSERM UMR_S910, 27 boulevard J. Moulin, 13005 Marseille, France, ²APHM, Laboratoire d'Anatomopathologie, ³APHM, Service de Neurochirurgie, ⁴APHM, Service de Neurologie, ⁵APHM, Laboratoire de génétique médicale, Hôpital de la Timone, 264 Rue Saint-Pierre, 13385 Marseille, France, ⁶UM76–UPMC, Inserm U974, CNRS UMR7215, 47, boulevard de l'Hôpital, La Pitié Salpêtrière–Bâtiment Babinski, 75651 Paris Cedex 13, France and ⁷Centre de Référence des Maladies Neuromusculaires, Nice University Hospital, 06000 Nice, France

Received April 8, 2013; Revised and Accepted June 6, 2013

Facio-scapulo-humeral dystrophy (FSHD) results from deletions in the subtelomeric macrosatellite D4Z4 array on the 4q35 region. Upregulation of the DUX4 retrogene from the last D4Z4 repeated unit is thought to underlie FSHD pathophysiology. However, no one knows what triggers muscle defect and when alteration arises. To gain further insights into the molecular mechanisms of the disease, we evaluated at the molecular level, the perturbation linked to the FSHD genotype with no a priori on disease onset, severity or penetrance and prior to any infiltration by fibrotic or adipose tissue in biopsies from fetuses carrying a short pathogenic D4Z4 array (n = 6) compared with fetuses with a non-pathogenic D4Z4 array (n = 21). By measuring expression of several muscle-specific markers and 4q35 genes including the DUX4 retrogene by an RT-PCR and western blotting, we observed a global dysregulation of genes involved in myogenesis including MYOD1 in samples with <11 D4Z4. The DUX4-fl pathogenic transcript was detected in FSHD biopsies but also in controls. Importantly, in FSHD fetuses, we mainly detected the non-spliced DUX4-fl isoform. In addition, several other genes clustered at the 4q35 locus are upregulated in FSHD fetuses. Our study is the first to examine fetuses carrying an FSHD-linked genotype and reveals an extensive dysregulation of several muscle-specific and 4q35 genes at early development stage at a distance from any muscle defect. Overall, our work suggests that even if FSHD is an adult-onset muscular dystrophy, the disease might also involve early molecular defects arising during myogenesis or early differentiation.

INTRODUCTION

Facio-scapulo-humeral dystrophy (FSHD) is an autosomal-dominant disorder, ranked as the most prevalent muscular dystrophy with an incidence of 7 of 100 000 (<http://www.orpha.net>).

Symptoms usually arise between the age of 20–40. Clinically, the disease manifests predominantly as facial weakness with progression to the upper body and then to the lower extremities with a characteristic asymmetric involvement of certain groups of skeletal muscles. There is a marked inter-

and intra-familial heterogeneity in FSHD clinical expression, which probably depends on multiple genetic and environmental factors unknown yet. Importantly, FSHD is linked to deletion of an integral number of a 3.3 kb tandem macrosatellite repeat arranged as a head-to-tail array in the subtelomeric 4q35 region. Typically, non-affected individuals carry between 11–100 copies of this D4Z4 element while patients with FSHD present between 1–10 units. Distal to D4Z4, a region defines two allelic forms, 4qA and 4qB. The 4qA sequence is characterized by the presence of an array of the 68 bp β -satellite repetitive DNA and the pLAM sequence abutting the last D4Z4 repeat and containing a polyadenylation site (1,2). Both qA and qB alleles are equally common in the population but FSHD is associated with the 4qA allele (3). In unaffected individuals, long D4Z4 arrays harbor heterochromatin features, whereas FSHD-linked short D4Z4 arrays contain epigenetic marks of non-transcribed euchromatin (4–6). D4Z4, which contains ~70% of CpG sites is hypermethylated at the DNA level in normal cells but hypomethylated in both 4q-linked (FSHD1) and phenotypic (FSHD2) FSHD.

So far, the search for a FSHD candidate gene has been controversial. The expression of several 4q35 genes has been determined in biopsies and primary myoblasts derived from patients, but their involvement and dysregulation in FSHD remain unclear with some authors observing a transcriptional downregulation of several of them (7–10) while other do not (11–13). Among the candidate genes for FSHD, the DUX4 retrogene localized within D4Z4 is suspected to play a key role (1,2,14,15). Each D4Z4 contains in its distal part an open reading frame encoding a putative protein containing two homeoboxes (1). In individuals carrying the 4qA allele downstream of the repeat, this DUX4 sequence can be transcribed through the last D4Z4 repeat and the proximal 4qA sequence leading to the production of a pre-messenger RNA with introns and a polyadenylation site (2,14,15). This pre-messenger RNA can be spliced into at least three different mRNAs of different sizes. In patient's samples, full-length DUX4 transcripts (DUX4-fl) are detected at a very low level (in 1 of 1000 muscle cell nuclei) (15). In the current model, DUX4-fl might encode a toxic double homeobox transcription factor responsible for the activation of ~500 other genes in skeletal muscle (16–18). Still, a major issue with the DUX4 hypothesis is its extremely low abundance, the difficulty to detect the protein in patient's biopsies limiting its use as a biomarker for diagnosis or prognosis and the time window in which DUX4 becomes pathogenic.

Thus, the search for genes dysregulated in FSHD led so far to controversial conclusions and different reasons might explain these discrepancies. Furthermore, if the genetic defect is present constitutively in most of the FSHD cases, the clinical signs of the disease only appear late in life and nothing is known on the molecular features of the presymptomatic muscle, in particular during development. Therefore, we investigated whether biopsies from fetuses carrying a short D4Z4 array display a developmental-specific phenotype compared with normal samples and quantified the expression level of different muscle-specific and 4q35 genes, including DUX4. Collectively, our data reveal changes in expression of several muscle-specific genes in fetuses with a short D4Z4 array compared with normal specimen together with the dysregulation of several 4q35 genes, suggesting that muscle homeostasis is affected as early as the fetal stage in FSHD.

RESULTS AND DISCUSSION

Key questions with regard to FSHD is how and when the muscular phenotype arises, the identity of the gene(s) involved in the pathology and the link between D4Z4 and the regulation of the FSHD-causing gene(s) in specific muscles. Therefore, to evaluate at the molecular level the perturbation linked to the FSHD genotype with no a priori on disease onset, severity or penetrance and prior to any infiltration by fibrotic or adipose tissue, we explored the expression level of several muscle-specific genes during the second trimester of gestation in six fetuses carrying a short pathogenic D4Z4 array (FSHD carriers) compared with fetuses with a long non-pathogenic D4Z4 array (controls, $n = 21$) at different gestational ages by measuring the expression of (i) several muscle-specific genes, (ii) DUX4 and (iii) different 4q35 genes.

During the human intrauterine development, multinucleated primary myotubes appear around the 5th week of gestation, early muscle fibers around the 11th week while after the 20th week, most muscle fibers are packed with myofibrils and display peripheral nuclei similar to adult muscle. In fetuses, muscle mass increases predominantly by proliferative growth of myoblasts. Between the 20th and 24th weeks of gestation, innervation enhances muscle development and differentiation. Up to the 25th week, muscle displays a hyperplastic stage with an increase in cell number. Then, the cell size and muscle mass increase rapidly by hypertrophy while the postnatal growth of muscle is mostly characterized by remodeling of pre-existing fibers (19). Two types of muscle fibers are present during the fetal period, but the distinction cannot be made until the 18–20th week of gestation.

With regard to the developmental pattern of muscle-specific genes, our results in quadriceps indicate an increase in the expression of several genes encoding structural proteins such as sarcoglycan (7.7-fold), calpain 3 (4.4-fold), dysferlin (3.1-fold). Also upregulation of several genes encoding the myosin heavy chain was evidenced, either for genes corresponding to type I, slow oxidative red fibers (MYH1, 8.25-fold; MYH7, 3.75-fold) or type II, fast oxidative red fibers, MYH2 (56-fold) in FSHD fetuses (Fig. 1A; Supplementary Material, Table S3; at least 2-fold change in expression; P -value ≥ 0.05). Furthermore, a significant upregulation of genes encoding proteins involved in muscle remodeling (MURF1 (TRIM63), 5.3-fold) and calcium release was also observed at the RNA (RYR1-1, 10.7-fold; CACNA1S, 7.7-fold) and protein level (RYR1, 2.8-fold, Fig. 1B). Concerning transcription factors involved in muscle differentiation, both MYOD1 and MRF4 are upregulated in FSHD fetuses (6.8- and 21.4-fold, respectively, Fig. 1A; Supplementary Material, Table S3).

Using the same criteria, expression analysis was done on adult quadriceps biopsies from six non-carrier individuals and seven FSHD patients. Among the different genes mentioned above, only MYOD was found to be significantly upregulated (3.8-fold) in FSHD as described elsewhere (13) (Supplementary Material, Table S3, and Fig. S1). Interestingly, MMP9 expression usually modulated in dystrophic muscle is not modified in FSHD fetuses (Fig. 1A), but seems decreased in adult FSHD patients (Supplementary Material, Fig. S1).

The development and plasticity of striated muscle are due to finely tuned networks acting at different levels from the prenatal

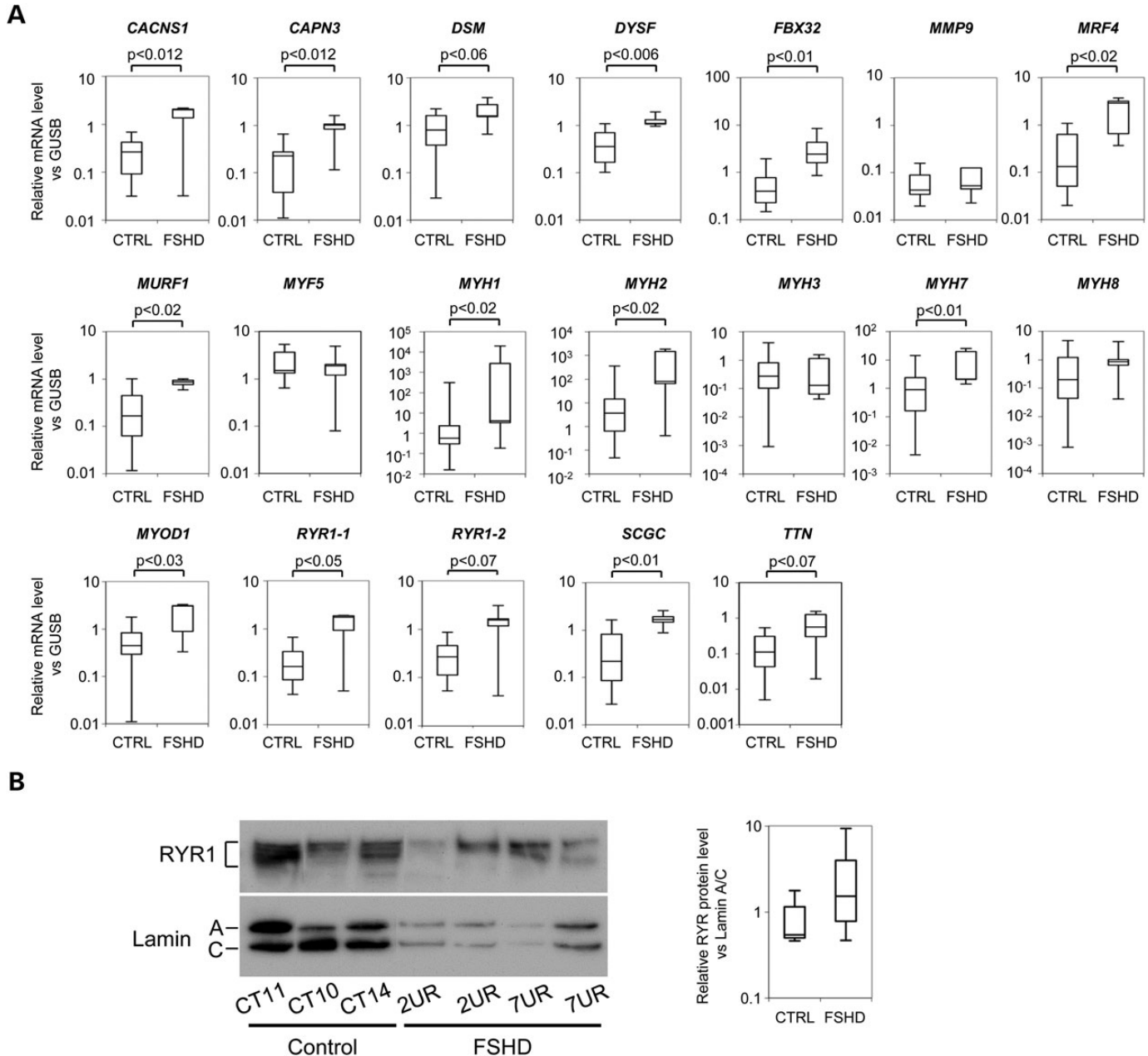


Figure 1. Transcriptional analysis of muscle-specific genes in fetal and adult muscle. (A) The box plots correspond to GUSB normalized data. P-values are indicated in each box. Experiments were done in triplicate. Horizontal lines are medians, the boxes correspond to the 50th percentile and the whiskers to the 75th percentile. Normalized transcription and P-values are given in Supplementary Material, Table S2. (B). Immunodetection of the RYR1 protein by western blotting (upper panel). Detection of Lamins A/C western was used as a reference (lower panel). The left panel corresponds to the quantification by densitometry of the RYR1/Lamins A/C signals. Boxes correspond to the 25th percentile and the whiskers to the 75th percentile. Horizontal lines correspond to the median.

to the post-natal period. Aberrant transcription of genes involved in the proliferation or differentiation of muscle cells is usually associated with intrinsic muscle defect. Histological observation of affected muscle from FSHD patients did not reveal any specific pathological change except, in some cases, changes in the fiber size, presence of angular fibers, infiltration by adipocytes, replacement of muscle fibers by fibrotic and inflammatory tissues or subtle sarcolemmal reorganization, which might change force generation by type II fibers (20,21). Also, from earlier work on primary cultures, necrotic features have been

described for FSHD myoblasts, which are more sensitive to stress, while FSHD myocytes fuse at a faster rate compared with controls but retain the appearance of an undifferentiated state. We report here, for the first time, global molecular changes in the fetal myogenic differentiation program, prior to any sign of dystrophic phenotype, suggesting that subtle molecular changes precede the dystrophic phenotype. Our work suggests that the clinical FSHD-specific phenotype might find its origin at very early developmental stages prior to any clinical sign of the disease. As suggested by others, even if FSHD is an

adult-onset muscular dystrophy, the disease might also depend on defects during myogenesis or early differentiation. Type-I slow-twitch oxidative fibers are preferentially found at early developmental stages, while type-II (fast twitch glycolytic) increases after the 26th week (22,23). Then, developmental myosin isoforms are progressively replaced by the adult isozymes. The marked increase in MYH2 in FSHD fetuses compared with controls suggests either modification in fiber conversion as observed in adults affected with FSHD (20,24) or a change in the timing of replacement since the change from type I to type II is usually seen after the 26th week of gestation. Furthermore, in FSHD, the presence of immature fibers expressing fetal myosin has been observed (25,26), suggesting also that maintenance or reactivation in the post-natal or adult muscle of factors acting during the intrauterine period might contribute to the cascade of alterations leading to the pathology.

The identity of the gene(s) involved in the pathology remains a matter of debate, but ectopic expression of the DUX4 retrogene from the last D4Z4 repeat and the distal qA region has been proposed as the main cause of the FSHD phenotype (1,2,15,16). To address whether DUX4 was present at early developmental stages, we used oligodT primed complementary DNA (cDNA) and PCR amplification with the primers previously described (15) to determine the presence of the full-length DUX4 transcript (DUX4-fl), blindly to the genetic background, in different fetal muscles from 5 FSHD fetuses and 21 controls (Fig. 2A, Table 1), but also in other tissues normally not affected in FSHD (diaphragm, spinal cord, skin, brain, and kidney). We detected the DUX4-fl mRNA in four of five FSHD samples either in biceps (one of four tested) or quadriceps (four of four) (Table 1), but also in non-affected somatic tissues such as skin, brain or kidney (Supplementary Material, Table S4), confirming that DUX4 transcription can be observed in tissues (muscular and non-muscular) from individuals carrying a shortened D4Z4 array as observed in muscular biopsies from adult FSHD patients (15). Moreover, as observed in healthy adult muscle (27), we also report the presence of the DUX4-fl in fetal biopsies from genetically unaffected fetuses although at a lower frequency (in 5 out of 17 biceps and 5 out of 21 quadriceps tested, Table 1), but also in non-muscular samples, not affected in the disease (Supplementary Material, Table S4). In adult samples, we confirmed the presence of the DUX4-fl in quadriceps of both controls (three of six) and FSHD patients (three of seven).

In FSHD patients carrying the 4qA allele downstream of the repeat, the transcription of DUX4 through D4Z4 and the proximal 4qA sequence produces a pre-messenger RNA with introns (1,2,15) spliced into different mRNAs varying in size and containing a polyadenylation site. A short DUX4 transcript (DUX4-s) can be detected in myoblasts from patients and 50% of controls. The long DUX4 transcript (DUX4-fl) is detected at a very low level (in 1 of 1000 muscle nuclei) in 50% of FSHD patient's biopsies, 40% expressing the DUX4-s and the others 10% presenting no expression of either form (15), but also in primary myoblasts and muscle biopsies from non-carrier unaffected adults (27).

In order to confirm specificity and sequence, we sequenced all PCR products amplified with the DUX4-fl primers and observed striking differences in splicing between FSHD carriers and controls. In most of the DUX4-positive samples from FSHD carriers (four out of five), we amplified the non-spliced DUX4-fl isoform,

while eight out of nine controls express the spliced DUX4-fl. Indeed, two alternative forms have been described for the DUX4-fl pathogenic transcripts differing by alternative splicing of a 135 bp-long intron in the 3' UTR (Fig. 2B). The termination codon of the DUX4 protein is located before the first alternative intron, and splicing is not expected to modify the frame or sequence of the pathogenic DUX4 protein.

Our data on DUX4 are consistent with the recent report showing the prevalence of a long DUX4-fl transcript in approximately half of the FSHD biopsies tested (15), but also in unaffected individuals (27). However, we extend here these findings by demonstrating that (i) DUX4 is expressed as early as fetal development, together with DUX4 targets (Ferreboeuf et al., submitted), in muscular and also non-muscular somatic tissues; (ii) both spliced and non-spliced DUX4-fl isoforms are found in adult muscle (iii) in fetal FSHD samples, RNA processing of the first intron is rarely observed, suggesting a FSHD-specific control of splicing. However, after their initial description (2,14,15), the respective regulation, the presence or role of the two variants has not been further investigated or discussed, and recent reports focusing on the biological function of the DUX4 protein describe a DUX4-fl encoding vector without any precision on whether it corresponds to the spliced or non-spliced form. After cloning of the PCR products and sequencing, we did not observe any co-expression of the spliced or non-spliced DUX4-fl suggesting that, as observed for DUX4-s and DUX4-fl, the two isoforms might be mutually exclusive. Nevertheless, the underlying mechanisms regulating alternative splicing and their respective role remain to be investigated.

Based on the hypothesis of a position effect mechanism, gradual changes in the expression of genes located at the 4q35 locus have been proposed for FSHD. In this model, relaxation and long-range chromatin-dependent mechanisms might be associated with increased and variable expression of one to several genes at distance of the D4Z4 array. So far, upregulation of the 4q35 genes has been a matter of debate with some authors showing an increase in expression of several of them (8–10) and others disputing these findings (11–13) in the absence of a gradient correlating distance from D4Z4 and gene expression (11,12,28). The reasons for such discrepancies are not known, but might depend on the type of sample (biopsies versus primary myoblasts or myotubes), type of muscle but also inter-individual variability linked to genetic or epigenetic factors. To limit the impact of such variability and elucidate the regulation of the 4q35 genes in individuals carrying a short D4Z4 array, we selected 36 genes located in the 4q35 region and measured their expression at the RNA level in muscular biopsies from 6 FSHD and 21 control fetuses. For at least 14 of the 4q35 genes, a significant increase in expression was observed in quadriceps (Fig. 2C; Supplementary Material, Table S4; at least 2-fold change in expression; P -value ≥ 0.05). We were not able to detect FRG2 transcripts in fetal biopsies, but observed FRG1 (10.9-fold; $P \geq 0.04$) and SLC25A4 (ANT1, 6-fold, $P \geq 0.005$) upregulation. In addition, the PDLIM3 gene (2.17-fold) involved in skeletal muscle function is also upregulated in FSHD fetuses. In adults, upregulation of ANT1 has already been reported (8,29) and proposed as an early event in the pathological process (9), while no noticeable changes were reported for PDLIM3 (ALP) (12,30) involved in Z-disc organization and regulation of the myogenic differentiation transcriptional network (31). Distal to

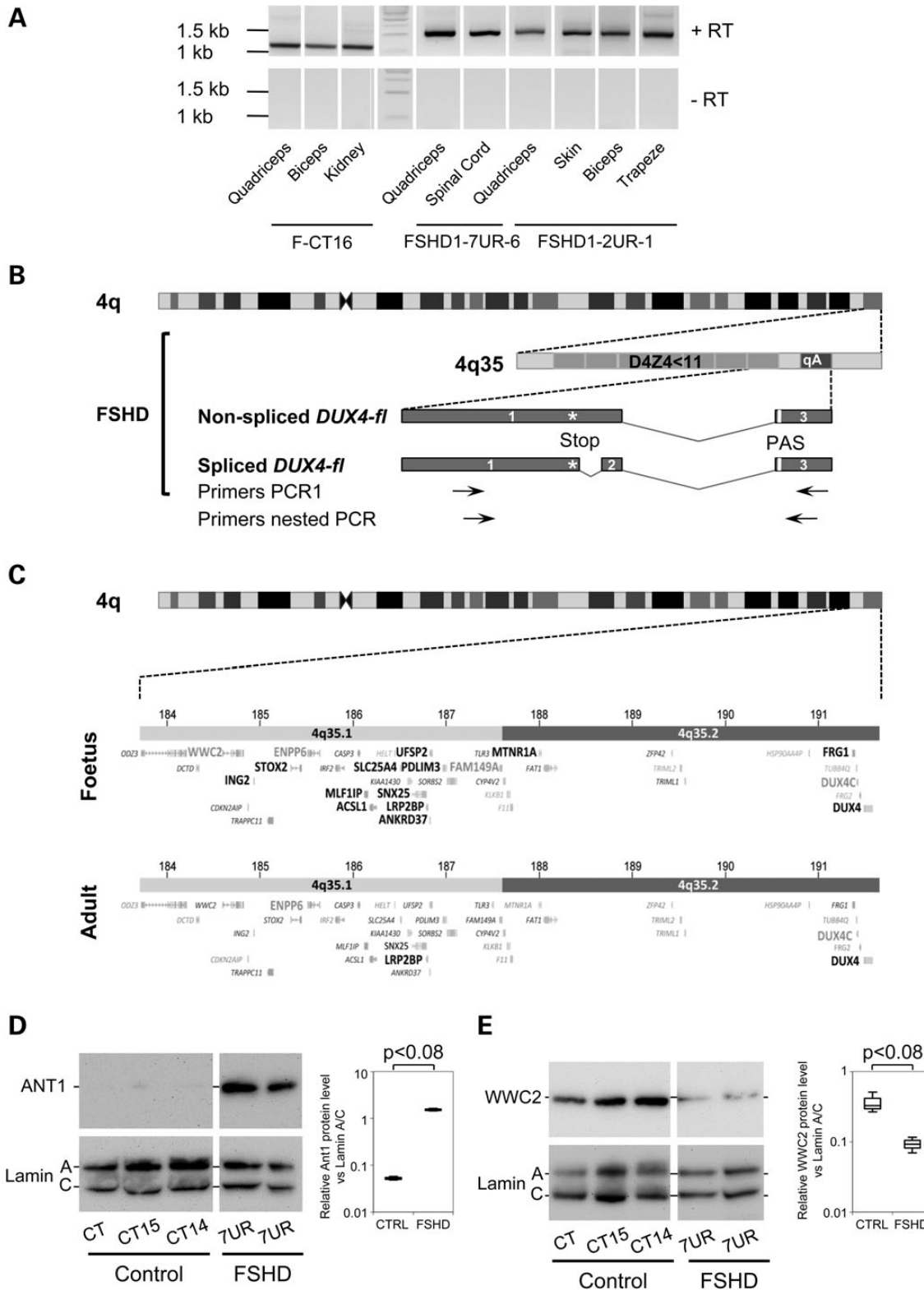


Figure 2. Expression of the 4q35 genes in fetal samples. (A). Upper panel, gel showing PCR products obtained after reverse transcription and nested PCR amplification on RNA extracted from muscle biopsies (biceps and quadriceps) and other somatic tissues of fetuses carrying a short D4Z4 array (FSHD1-7UR-6; FSHD1-2UR-1) or non-carriers (F-CT16). Lower panel corresponds to amplification from samples not treated with reverse transcriptase (RT) (B). Schematic representation of the last D4Z4 repeat encoding the two *DUX4-fl* isoforms. The *DUX4* exons indicated have grey boxes and numbered (1–3). The position of the polyadenylation site (PAS), stop codon (star) and primers used for RT–PCR are indicated. (C). Diagram showing changes in gene expression across the 4q35 region in quadriceps from fetuses (n ¼ 6; upper panel) and adults (n ¼ 7; lower panel) carrying a shortened D4Z4 array compared with controls (fetuses, n ¼ 21; adults, n ¼ 6). Black and upper cases: upregulated genes; Grey upper case: downregulated genes; black: no significant change; grey: not expressed. Experiments were done in triplicate. Normalized

Table 1. Amplification of the DUX4-fl isoform in fetal and adult muscles

	Gestational age (weeks)	Biceps	Quad.
Fetuses			
FSHD1-2UR-1	16	(+) ²	(+) ²
FSHD1-4UR-2	16	–	(+) ²
FSHD1-2UR-4	25	N.A.	(+) ¹
FSHD1-7UR-5	26	–	N.A.
FSHD1-7UR-6	26	–	(+) ²
F-CT1	12	–	–
F-CT2	13	–	–
F-CT3	14	–	(+) ¹
F-CT4	14	(+) ¹	–
F-CT5	14	–	–
F-CT6	14	–	–
F-CT7	15	–	(+) ¹
F-CT8	18	N.A.	(+) ¹
F-CT9	18	(+) ¹	–
F-CT10	20	–	–
F-CT11	21	N.A.	–
F-CT12	23	–	–
F-CT13	24	N.A.	–
F-CT14	25	–	(+) ¹
F-CT15	25	(+) ¹	–
F-CT16	26	(+) ²	(+) ²
F-CT17	34	N.A.	–
F-CT18	37	(+) ¹	–
F-CT19	37	–	–
F-CT20	37	–	–
F-CT21	37	–	–
Adults			
A-CT1	51	–	N.A.
A-CT2	77	–	N.A.
A-CT3	69	N.A.	(+) ¹
A-CT4	45	N.A.	(+) ¹
A-CT5	49	N.A.	(+) ¹
A-CT6	52	N.A.	–
FSHD1-1	54	N.A.	(+) ²
FSHD1-2	66	N.A.	–
FSHD1-3	61	N.A.	–
FSHD1-4	61	N.A.	–
FSHD1-5	57	N.A.	–
FSHD1-6	57	N.A.	(+) ¹
FSHD1-7	45	N.A.	(–) ²

The table recapitulates the results of at least three independent experiments. Samples positive in at least two out of the three RT–PCR assays have been sequenced. (2) indicates that the DUX4-fl transcript was not detected. (+) corresponds to samples where DUX4-fl was detected (either biceps or quadriceps). Each amplicon was sequenced and presence of either the spliced (+)¹ or non-spliced (+)² DUX4-fl isoform is indicated.

PDLIM3, the SORBS2 gene regulated by the PDLIM3-SORBS2 distal enhancer (32) during muscle development is not differentially expressed in FSHD (Fig. 2C, Supplementary Material, Fig. S4). In the vicinity of PDLIM3A and SCL25A4, four other genes are upregulated (LRP2BP, ACSL1, ANKRD37 and UFSP2). LRP2BP (6.11-fold) regulates LRP2 function and HDL endocytosis and ACSL1 (4.3-fold) encodes the acyl-coenzyme A synthetase long chain 1 implicated in the adaptation

of several tissues including muscle to aerobic metabolism. The function of ANKRD37 is not known, but might be activated by HIF1 in response to hypoxia while UFSP2 encodes an ubiquitin-fold modifier protease. Their respective function in muscle remains obscure. Two additional genes are also upregulated in FSHD: MNTR1A (5.43-fold) encoding a melatonin receptor and ING2 (2.12-fold) encoding a chromatin-remodeling factor implicated in muscle differentiation. Beside DUX4C (7.3-fold, $P > 0.05$), two genes ENPP6 (ectonucleotide pyrophosphatase/phosphodiesterase 6), WWC2 (WW and C2 domain containing protein) appear to be slightly down-regulated in FSHD compared with control muscles (2.3-; -1.9- respectively, $P < 0.05$; Figure 2C; Supplementary Material, Table S5, Supplementary Material, Fig. S2C). SCL25A4 and WWC2 were immunodetected by western blotting and densitometry analysis confirmed their respective up- (29.6-fold, $P > 0.1$; Fig. 2D) or down-regulation (3.5-fold; $P > 0.1$; Fig. 2E) in fetal quadriceps FSHD samples compared with age-matched controls. As a comparison, we examined the expression status of the 36 genes in seven FSHD adult biopsies and six controls (Supplementary Material, Fig. S2 and Table S5). In agreement with the literature but in contrast to fetuses, we did not observe any significant dysregulation for these genes in adult muscles.

In fetuses, shortening of the D4Z4 array is accompanied with dysregulation in a set of genes involved in myogenesis suggesting an altered myogenesis as early as the fetal stage. The genetic defect linked to FSHD might directly or indirectly disrupt the myogenic program as discussed before (33). Wasting and weakening in adult FSHD muscle might thus result from the long-term accumulation of defects such as structural changes, fiber conversion or response to stress arising as early as the fetal life. In FSHD, the rationale for the selective and asymmetrical involvement of specific groups of muscle remains unclear, since all the affected muscle do not originate from the same somite or same region in the embryo. However, one can speculate that common regulatory pathways regulating migration or differentiation of muscle cells are affected during muscle differentiation, and that the abnormality leading to the pathology begins during development. Strikingly, many of the 4q35 genes upregulated in muscles of FSHD carriers are also implicated at different levels in muscle function and homeostasis. Thus, dysregulation of the 4q35 locus together with changes in muscle-specific genes might alter the capacity of the muscle fiber to respond to stress during adult life leading, in turn, to the progressive weakening observed at the clinical level. We did not find any evidence of a common regulatory pathway for the different genes upregulated in FSHD fetuses (Supplementary Material, Figure S3) but, interestingly, many of the genes strongly upregulated in fetuses (ACSL1, LRP2BP, PDLIM3, SCL25A4, UFSP2 and MNTR1A) are clustered in a region located 5–5.5 Mb upstream of the most proximal D4Z4 element, flanked by two genomic regions located to the nuclear periphery (lamin attachment domains). This observation, in light of our previous work (4,34,35) and data from others (36,37), suggests that

transcription values are indicated in Supplementary Material, Table S6. We also observed upregulation of PDLIM3, UFSP2, STOX2, ING2, ANTI in biceps from FSHD fetuses. The corresponding box plots are shown in Supplementary Material, Figure S3. Upregulation is not correlated either with orientation of the gene or with the presence of a CpG island at the promoter region. (D) Western blot of ANTI/SLC25A4 and (E) WWC2 protein. Lamins A/C were used as a reference. On the left of each gel, quantification by densitometry of the protein-specific band/Lamins A/C signals. Horizontal lines are medians, the boxes correspond to the 25th percentile and the whiskers to the 75th percentile.

conformational and topological changes mediated by D4Z4 array shortening might modify the regulation of the whole 4q35 region. Furthermore, with regard to expression of the 4q35 genes and muscle-specific genes, our results suggest that FSHD might not simply result from overexpression of a single gene or retrogene, but more likely from a cascade of dysregulations at different stages of the muscle development and maturation associated with the DUX4-dependent cascade or on other D4Z4-dependent modification.

At this step, it remains difficult to connect into a single model the mechanism of the disease and the respective contribution of the different DUX4-fl isoforms and other genes in the pathogenesis of this complex muscular dystrophy. Transcriptional mechanisms governing D4Z4 transcription are likely to be important for understanding the pathology since a large number of small, polyadenylated or long non-coding RNA emanating from the repeat have been described (14,38,39). In general, our data suggest that reduction in the number of D4Z4 might lead to the pathology by enhancing in cis expression of different genes including DUX4. However, the absence of the pathogenic DUX4-fl in a number of FSHD cases (15), its presence non-muscular somatic tissues in FSHD carriers but also in unaffected individuals (27) suggest that other alterations or pathways might contribute to the muscular phenotype as a direct or indirect consequence of this cis effect. FSHD is a highly variable disease with a high variability in penetrance within families even in individuals carrying the same number of repeats or monozygotic twins (25,40,41), suggesting that environment and lifestyle modify expression of the gene(s) involved in the disease. Overall, our work does not exclude a role for a stable DUX4 transcript linked to a short or relaxed D4Z4 array and the presence of a distal functional polyadenylation site, but suggests that DUX4 detection is not the only marker of FSHD. Disease onset and progression might be associated with either dosage of the pathogenic transcript, stability or maintenance in the muscle fiber or burst of activation at critical stages. Of note, we were able to detect DUX4 as early as the 14th week of gestation in normal fetuses and 16th week in FSHD samples, corresponding to the formation of fibers and myofibrils (11th–20th week). Recent data indicate that DUX4 is regulated during embryogenesis (42), and that muscle is sensitive to DUX4 dosage at very early stages (43). Hence, the window of time in which DUX4 is activated might be critical for FSHD. However, dysregulation of different 4q35 and muscle-specific genes in fetal muscles also suggests that a cascade of events probably precedes FSHD symptoms, and that accumulation of defects at different stages of muscle development might contribute to the pathomechanisms.

MATERIALS AND METHODS

Biological samples

Fetal biopsies from FSHD ($n = 6$) and non-carrier fetuses ($n = 21$) were obtained after therapeutic abortion. Research was approved by the biomedicine agency (PFS13-006). The parents have provided written informed consent for the use of biopsies for medical research in accordance with the Declaration of Helsinki. Controls are neither carrier of any known genetic mutation nor affected by a muscular pathology. Adults have provided informed consent and muscles of clinically affected and

control individuals were obtained using a standardized muscle biopsy protocol. Controls were selected in the same age range and sex representation.

Quantitative RT-PCR

Total RNA was extracted from biopsies using the classical Chomczynski and Sacchi method (44). Reverse transcription of 1 mg of total RNA was performed using the Superscript II kit and oligo dT following manufacturer's instructions at 42°C for 50 min followed by inactivation at 70°C for 15 min (Life Technologies). Primers were designed using Primer Blast and Primer 3 (Supplementary Material, Table S1). Real-time PCR amplification was performed on a LightCycler 480 (Roche) using the SYBR green master mix. All PCRs were performed using a standardized protocol, and data were analyzed with the Lightcycler 480 software version 1.5.0.39 (Roche). Primer efficiency was determined by absolute quantification using a standard curve. For each sample, fold change was obtained by comparative quantification and normalization to expression of the GUSB standard gene. Similar results were obtained after normalization to β -actin or β 2M. For statistical analysis, data are expressed as means \pm SEM. Statistical significance was assessed by the non-parametric Wilcoxon–Mann–Whitney statistical test using the Gnumeric spreadsheet version 1.10.16.

RT-PCR for DUX4

One microgram of total RNA was used for first-strand cDNA synthesis using SuperScript II reverse transcriptase and oligo dT as described above. Primary PCRs were performed with Taq DNA polymerase (Euromedex) using 7% of the first-strand reaction as template in a total reaction volume of 30 μ l. Nested PCRs were performed on 1 μ l of the primary reaction. For DUX4-fl detection, the following primers were used: PCR1-forward: 5'-CCC CGA GCC AAA GCG AGG CCC TGC GAG CCT-3'; PCR1-reverse: 5'-GTA ACT CTA ATC CAG GTT TGC CTA GA-3'; PCR2-forward: 5'-CGG CCC TGG CCC GGG AGA ACG CGG CCC GC-3'; PCR2-reverse: 5'-TCT AAT CCA GGT TTG CCT AGAV CAG C-3' as described in (15). PCR products were examined on 3% Molecular Biology Grade Agarose gels stained with ethidium bromide. In order to discard any risk of contamination by genomic DNA, each RNA sample was treated with DNase I and PCR amplification was performed on RNA samples incubated in the absence of reverse transcriptase (RT-) or without addition of cDNA. In all cases, amplification products were only observed when RNA samples were reverse-transcribed. All PCR products were sequenced.

Western blot

Whole protein extracts were obtained from biopsies disrupted in 200 μ l extraction buffer (Tris-HCl pH8 100 mM, 10% SDS, 10 mM EDTA, 10% glycerol, protease inhibitor) using a T18 Ultra Turax. Proteins were separated by electrophoresis and transferred onto a PVDF membrane following the protocol recommended by the supplier for the Life Technologies NuPAGE system (including MOPS running buffers, Bis/Tris 4-12% gels and NuPAGE transfer buffer). PVDF membranes

were blocked for 1 h in 5% (w/v) non-fat dry milk in PBS-Tween (0.1% Tween-20 in PBS) and incubated for 90 min with the following primary antibodies: Lamin A/C (1 of 10 000, Clone 4C11, Sigma-Aldrich), WWC2 (1 of 1000, Sigma-Aldrich), SLC25A4 (1 of 1000, Sigma-Aldrich), RYR1 (1 of 1000, ref: ab2868, Abcam). After four washes in PBS-T, an anti-mouse IgG secondary antibody coupled to horseradish peroxidase (ThermoFisher) was incubated for 90 min (1/20 000). After washes, the signal was revealed by enhanced chemiluminescence (ECL, SuperSignal West Pico, Pierce) using a biospectrum imaging system (UVP).

SUPPLEMENTARY MATERIAL

Supplementary Material is available at HMG online.

ACKNOWLEDGEMENTS

We are indebted to all families for participating in this study. We wish to thank Mr Armand Tasmadjian for technical help. We also thank Dr Sabrina Sacconi for providing adult biopsies and critical reading.

Conflict of Interest statement. N.B. is the recipient of a fellowship from AFM (Association Française contre les Myopathies). S.R. is the recipient of a fellowship from ANR (FSHDecrypt, ANR-09-GENO-038), M.-C.G. is the recipient of a fellowship from the FSH Society and M.F. is the recipient of a fellowship from FSHD Global (Australia).

FUNDING

This study was funded by FSHD Global (Australia), the FSH Society (USA) and an ANR grant (FSHDecrypt, ANR-09-GENO-038) (to F.M.).

REFERENCES

- Gabriels, J., Beckers, M.C., Ding, H., De Vriese, A., Plaisance, S., van der Maarel, S.M., Padberg, G.W., Frants, R.R., Hewitt, J.E., Collen, D. et al. (1999) Nucleotide sequence of the partially deleted D4Z4 locus in a patient with FSHD identifies a putative gene within each 3.3 kb element. *Gene*, 236, 25–32.
- Lemmers, R.J., van der Vliet, P.J., Klooster, R., Sacconi, S., Camano, P., Dauwerse, J.G., Snider, L., Straasheijm, K.R., van Ommen, G.J., Padberg, G.W. et al. (2010) A unifying genetic model for facioscapulohumeral muscular dystrophy. *Science*, 329, 1650–1653.
- Lemmers, R.J., Wohlgenuth, M., Frants, R.R., Padberg, G.W., Morava, E. and van der Maarel, S.M. (2004) Contractions of D4Z4 on 4qB subtelomeres do not cause facioscapulohumeral muscular dystrophy. *Am. J. Hum. Genet.*, 75, 1124–1130.
- Ottaviani, A., Rival-Gervier, S., Boussouar, A., Foerster, A.M., Rondier, D., Sacconi, S., Desnuelle, C., Gilson, E. and Magdinier, F. (2009) The D4Z4 macrosatellite repeat acts as a CTCF and A-type lamins-dependent insulator in facio-scapulo-humeral dystrophy. *PLoS Genet.*, 5, e1000394.
- Tsumagari, K., Qi, L., Jackson, K., Shao, C., Lacey, M., Sowden, J., Tawil, R., Vedanarayanan, V. and Ehrlich, M. (2008) Epigenetics of a tandem DNA repeat: chromatin DNaseI sensitivity and opposite methylation changes in cancers. *Nucleic Acids Res.*, 36, 2196–2207.
- Zeng, W., de Greef, J.C., Chen, Y.Y., Chien, R., Kong, X., Gregson, H.C., Winokur, S.T., Pyle, A., Robertson, K.D., Schmiesing, J.A. et al. (2009) Specific loss of histone H3 lysine 9 trimethylation and HP1 gamma/cohesin binding at D4Z4 repeats is associated with facioscapulohumeral dystrophy (FSHD). *PLoS Genet.*, 5, e1000559.
- Gabellini, D., D'Antona, G., Moggio, M., Prella, A., Zecca, C., Adami, R., Angeletti, B., Ciscato, P., Pellegrino, M.A., Bottinelli, R. et al. (2006) Facioscapulohumeral muscular dystrophy in mice overexpressing FRG1. *Nature*, 439, 973–977.
- Gabellini, D., Green, M.R. and Tupler, R. (2002) Inappropriate gene activation in FSHD: a repressor complex binds a chromosomal repeat deleted in dystrophic muscle. *Cell*, 110, 339–348.
- Laoudj-Chenivresse, D., Carnac, G., Bisbal, C., Hugon, G., Bouillot, S., Desnuelle, C., Vassetzky, Y. and Fernandez, A. (2005) Increased levels of adenine nucleotide translocator 1 protein and response to oxidative stress are early events in facioscapulohumeral muscular dystrophy muscle. *J. Mol. Med.*, 83, 216–224.
- Rijkers, T., Deidda, G., van Koningsbruggen, S., van Geel, M., Lemmers, R.J., van Deutekom, J.C., Figlewicz, D., Hewitt, J.E., Padberg, G.W., Frants, R.R. et al. (2004) FRG2, an FSHD candidate gene, is transcriptionally upregulated in differentiating primary myoblast cultures of FSHD patients. *J. Med. Genet.*, 41, 826–836.
- Jiang, G., Yang, F., van Overveld, P.G., Vedanarayanan, V., van der Maarel, S. and Ehrlich, M. (2003) Testing the position-effect variegation hypothesis for facioscapulohumeral muscular dystrophy by analysis of histone modification and gene expression in subtelomeric 4q. *Hum. Mol. Genet.*, 12, 2909–2921.
- Winokur, S.T., Chen, Y.W., Masny, P.S., Martin, J.H., Ehmsen, J.T., Tapscott, S.J., van der Maarel, S.M., Hayashi, Y. and Flanagan, K.M. (2003) Expression profiling of FSHD muscle supports a defect in specific stages of myogenic differentiation. *Hum. Mol. Genet.*, 12, 2895–2907.
- Cheli, S., Francois, S., Bodega, B., Ferrari, F., Tenedini, E., Roncaglia, E., Ferrari, S., Ginelli, E. and Meneveri, R. (2011) Expression profiling of FSHD-1 and FSHD-2 cells during myogenic differentiation evidences common and distinctive gene dysregulation patterns. *PLoS One*, 6, e20966.
- Snider, L., Asawachaicharn, A., Tyler, A.E., Geng, L.N., Petek, L.M., Maves, L., Miller, D.G., Lemmers, R.J., Winokur, S.T., Tawil, R. et al. (2009) RNA transcripts, miRNA-sized fragments and proteins produced from D4Z4 units: new candidates for the pathophysiology of facioscapulohumeral dystrophy. *Hum. Mol. Genet.*, 18, 2414–2430.
- Snider, L., Geng, L.N., Lemmers, R.J., Kyba, M., Ware, C.B., Nelson, A.M., Tawil, R., Filippova, G.N., van der Maarel, S.M., Tapscott, S.J. et al. (2010) Facioscapulohumeral dystrophy: incomplete suppression of a retrotransposed gene. *PLoS Genet.*, 6, e1001181.
- Vanderplanck, C., Anseau, E., Charron, S., Stricwant, N., Tassin, A., Laoudj-Chenivresse, D., Wilton, S.D., Coppee, F. and Belayew, A. (2011) The FSHD atrophic myotube phenotype is caused by DUX4 expression. *PLoS One*, 6, e26820.
- Wallace, L.M., Garwick, S.E., Mei, W., Belayew, A., Coppee, F., Ladner, K.J., Guttridge, D., Yang, J. and Harper, S.Q. (2011) DUX4, a candidate gene for facioscapulohumeral muscular dystrophy, causes p53-dependent myopathy in vivo. *Ann Neurol*, 69, 540–552.
- Kowalijow, V., Marcowycz, A., Anseau, E., Conde, C.B., Sauvage, S., Matteotti, C., Arias, C., Corona, E.D., Nunez, N.G., Leo, O. et al. (2007) The DUX4 gene at the FSHD1A locus encodes a pro-apoptotic protein. *Neuromuscul. Disord.*, 17, 611–623.
- Ahn, S., Riccio, A. and Ginty, D.D. (2000) Spatial considerations for stimulus-dependent transcription in neurons. *Annu. Rev. Physiol.*, 62, 803–823.
- Lassche, S., Stienen, G.J., Irving, T.C., van der Maarel, S.M., Voermans, N.C., Padberg, G.W., Granzier, H., van Engelen, B.G. and Ottenheijm, C.A. (2013) Sarcomeric dysfunction contributes to muscle weakness in facioscapulohumeral muscular dystrophy. *Neurology*, 80, 733–737.
- Reed, P., Porter, N.C., Strong, J., Pumphlin, D.W., Corse, A.M., Luther, P.W., Flanagan, K.M. and Bloch, R.J. (2006) Sarcolemmal reorganization in facioscapulohumeral muscular dystrophy. *Ann. Neurol.*, 59, 289–297.
- Barbet, J.P., Thornell, L.E. and Butler-Browne, G.S. (1991) Immunocytochemical characterisation of two generations of fibers during the development of the human quadriceps muscle. *Mech. Dev.*, 35, 3–11.
- Butler-Browne, G.S., Barbet, J.P. and Thornell, L.E. (1990) Myosin heavy and light chain expression during human skeletal muscle development and precocious muscle maturation induced by thyroid hormone. *Anat. Embryol. (Berl)*, 181, 513–522.
- Celegato, B., Capitanio, D., Pescatori, M., Romualdi, C., Pacchioni, B., Cagnin, S., Vigano, A., Colantoni, L., Begum, S., Ricci, E. et al. (2006) Parallel protein and transcript profiles of FSHD patient muscles correlate to the D4Z4 arrangement and reveal a common impairment of slow to fast fibre

- différenciation and a general deregulation of MyoD-dependent genes. *Proteomics*, 6, 5303–5321.
25. Tupler, R., Barbierato, L., Memmi, M., Sewry, C.A., De Grandis, D., Maraschio, P., Tiepolo, L. and Ferlini, A. (1998) Identical de novo mutation at the D4F104S1 locus in monozygotic male twins affected by facioscapulohumeral muscular dystrophy (FSHD) with different clinical expression. *J. Med. Genet.*, 35, 778–783.
 26. Rahimov, F., King, O.D., Leung, D.G., Bibat, G.M., Emerson, C.P. Jr., Kunkel, L.M. and Wagner, K.R. (2012) Transcriptional profiling in facioscapulohumeral muscular dystrophy to identify candidate biomarkers. *Proc Natl Acad Sci USA*, 109, 16234–16239.
 27. Jones, T.I., Chen, J.C., Rahimov, F., Homma, S., Arashiro, P., Beermann, M.L., King, O.D., Miller, J.B., Kunkel, L.M., Emerson, C.P. Jr. et al. (2012) Facioscapulohumeral muscular dystrophy family studies of DUX4 expression: evidence for disease modifiers and a quantitative model of pathogenesis. *Hum. Mol. Genet.*, 21, 4419–4430.
 28. Masny, P.S., Chan, O.Y., de Greef, J.C., Bengtsson, U., Ehrlich, M., Tawil, R., Lock, L.F., Hewitt, J.E., Stocksdales, J., Martin, J.H. et al. (2010) Analysis of allele-specific RNA transcription in FSHD by RNA-DNA FISH in single myonuclei. *Eur. J. Hum. Genet.*, 18, 448–456.
 29. Tsumagari, K., Chang, S.C., Lacey, M., Baribault, C., Chittur, S.V., Sowden, J., Tawil, R., Crawford, G.E. and Ehrlich, M. (2011) Gene expression during normal and FSHD myogenesis. *BMC Med Genomics*, 4, 67.
 30. Bouju, S., Pietu, G., Le Cunff, M., Cros, N., Malzac, P., Pellissier, J.F., Pons, F., Leger, J.J., Auffray, C. and Dechesne, C.A. (1999) Exclusion of muscle specific actinin-associated LIM protein (ALP) gene from 4q35 facioscapulohumeral muscular dystrophy (FSHD) candidate genes. *Neuromuscul. Disord.*, 9, 3–10.
 31. Pomies, P., Pashmforoush, M., Vegezzi, C., Chien, K.R., Auffray, C. and Beckerle, M.C. (2007) The cytoskeleton-associated PDZ-LIM protein, ALP, acts on serum response factor activity to regulate muscle differentiation. *Mol. Biol. Cell*, 18, 1723–1733.
 32. McCord, R.P., Zhou, V.W., Yuh, T. and Bulyk, M.L. (2011) Distant cis-regulatory elements in human skeletal muscle differentiation. *Genomics*, 98, 401–411.
 33. Winokur, S.T., Barrett, K., Martin, J.H., Forrester, J.R., Simon, M., Tawil, R., Chung, S.A., Masny, P.S. and Figlewicz, D.A. (2003) Facioscapulohumeral muscular dystrophy (FSHD) myoblasts demonstrate increased susceptibility to oxidative stress. *Neuromuscul. Disord.*, 13, 322–333.
 34. Ottaviani, A., Schluth-Bolard, C., Gilson, E. and Magdinier, F. (2010) D4Z4 as a prototype of CTCF and lamins-dependent insulator in human cells. *Nucleus*, 1, 30–36.
 35. Ottaviani, A., Schluth-Bolard, C., Rival-Gervier, S., Boussouar, A., Rondier, D., Foerster, A.M., Morere, J., Bauwens, S., Gazzo, S., Callet-Bauchu, E. et al. (2009) Identification of a perinuclear positioning element in human subtelomeres that requires A-type lamins and CTCF. *EMBO J.*, 28, 2428–2436.
 36. Bodega, B., Ramirez, G.D., Grasser, F., Cheli, S., Brunelli, S., Mora, M., Meneveri, R., Marozzi, A., Mueller, S., Battaglioli, E. et al. (2009) Remodeling of the chromatin structure of the facioscapulohumeral muscular dystrophy (FSHD) locus and upregulation of FSHD-related gene 1 (FRG1) expression during human myogenic differentiation. *BMC Biol.*, 7, 41.
 37. Pirozhkova, I., Petrov, A., Dmitriev, P., Laoudj, D., Lipinski, M. and Vassetzky, Y. (2008) A functional role for 4qA/B in the structural rearrangement of the 4q35 region and in the regulation of FRG1 and ANT1 in facioscapulohumeral dystrophy. *PLoS One*, 3, e3389.
 38. Cabcianca, D.S., Casa, V., Bodega, B., Xynos, A., Ginelli, E., Tanaka, Y. and Gabellini, D. (2012) A long ncRNA links copy number variation to a polycomb/trithorax epigenetic switch in FSHD muscular dystrophy. *Cell*, 149, 819–831.
 39. Block, G.J., Petek, L.M., Narayanan, D., Amell, A.M., Moore, J.M., Rabaia, N.A., Tyler, A., van der Maarel, S.M., Tawil, R., Filippova, G.N. et al. (2012) Asymmetric bidirectional transcription from the FSHD-causing D4Z4 array modulates DUX4 production. *PLoS One*, 7, e35532.
 40. Griggs, R.C., Tawil, R., McDermott, M., Forrester, J., Figlewicz, D. and Weiffenbach, B. (1995) Monozygotic twins with facioscapulohumeral dystrophy (FSHD): implications for genotype/phenotype correlation. *FSH-DY Group. Muscle Nerve Suppl.*, 2, S50–S55.
 41. Tawil, R., Storvick, D., Feasby, T.E., Weiffenbach, B. and Griggs, R.C. (1993) Extreme variability of expression in monozygotic twins with FSH muscular dystrophy. *Neurology*, 43, 345–348.
 42. Krom, Y.D., Thijssen, P.E., Young, J.M., den Hamer, B., Balog, J., Yao, Z., Maves, L., Snider, L., Knopp, P., Zammit, P.S. et al. (2013) Intrinsic epigenetic regulation of the D4Z4 macrosatellite repeat in a transgenic mouse model for FSHD. *PLoS Genet.*, 9, e1003415.
 43. Mitsuhashi, H., Mitsuhashi, S., Lynn-Jones, T., Kawahara, G. and Kunkel, L.M. (2013) Expression of DUX4 in zebrafish development recapitulates facioscapulohumeral muscular dystrophy. *Hum. Mol. Genet.*, 22, 568–577.
 44. Chomczynski, P. and Sacchi, N. (2006) The single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction: twenty-something years on. *Nat Protoc.*, 1, 581–585.